a folha

Boletim da língua portuguesa nas instituições europeias

http://ec.europa.eu/translation/portuguese/magazine

N.º 53 — primavera de 2017

Campeão

Agustín Jiménez
Antigo tradutor da Direção-Geral da Tradução — Comissão Europeia

[Versão original, em língua espanhola, publicada no n.º 151, janeiro-fevereiro de 2017, do boletim *puntoycoma*[1] dos tradutores espanhóis das instituições da União Europeia.]

A palavra portuguesa que ficou em segundo lugar na eleição da palavra do ano de 2016[2] foi «campeão». Ganhou «geringonça» (construção frágil, como a atual coligação governamental), mas os portugueses sabiam que eram os melhores no futebol, e a pretensão da palavra do ano é expressar modas linguísticas que se casem, além disso, com a atualidade[3]. Na Islândia, o país mais literário do mundo, a palavra de 2016 foi cunhada por um escritor («*hrútskýring*»: o equivalente do inglês «*mansplaining*»[4]); na Noruega, o primeiro-ministro usou-a num discurso «*hverdagsintegrering*» (integração quotidiana dos refugiados); em sítios supostamente provocadores como a Suécia, as palavras do ano são indicadas pela Academia; na Bélgica, os leitores do *Le Soir*[5] concederam uma menção honrosa a «*putaclic*»[6] (algo como «agarra-cliques») e os flamengos, respondendo à iniciativa da editora neerlandesa Van Dale, consagraram «*samsonseks*»[7] (ato sexual praticado enquanto as crianças veem televisão). Em Espanha, o júri da Fundéu optou pela facilidade e escolheu «*populismo*», deixando para segundo plano outras palavras como «*abstenciocracia*» ou «*ningufonear*»[8], para alívio da maioria, que nunca as ouvira. Os franceses escolheram uma obviedade, «*refugiés*», o que Portugal e Espanha já haviam feito em 2015, mas fizeram-no com pompa: na 12.ª edição do «Festival du mot»[9]. Em Itália, país de verbo caudaloso, as palavras vencedoras foram escolhidas, caudalosamente e por sondagem, pelos internautas[10].

Palavras/palavrões, que maioritariamente sugerem fobias comuns, palhaços que assustam, o jogo do Pokémon, a proibição do burquíni ou o americano que acabam de tornar presidente. Um termo bem situado na lista de finalistas gaulesa foi o aglutinador «*république*», de nulo interesse linguístico, mas assaz significativo, porque o segundo prémio foi para «*terrorisme*». Neste sentido introspetivo, os dinamarqueses foram mais explícitos. Perante a atualidade e as ameaças externas, o vocábulo

vencedor foi «*danskhed*» (algo como «dinamarquidade»). Mais bela é a introspeção inerente à palavra do ano consagrada pela Real Academia Galega. Cumprindo-se em 2016 o centenário das «irmandades da fala», que impulsionaram o amor à mãe Galiza e à língua galega no primeiro quarto do século passado, a palavra galega do ano foi «irmandade»(11).

Em terceiro lugar da lista de termos candidatos galegos figura «alburgada»(12) (embuste), o que parece mais interessante do que «pós-verdade», o decalque da palavra escolhida pelo dicionário Oxford(13). «*Post-truth*» é um termo documentado desde 1992, que, se no resto do mundo evoca o incómodo dos que veem preconceitos e boatos prevalecerem sobre os factos, no Reino Unido foi diretamente inspirado pelas circunstâncias do Brexit, palavra que na lista do Oxford vem acompanhada pelo seu derivado «*brexiteer*» (partidário da saída da União Europeia). Um sufixo contagioso, esse «*exit*», que também gerou «Grexit» e «Frexit» (*Lepen dixit*). Neste mesmo espírito, os alemães propuseram expressões como «*posfaktisch*»(14) e «Brexit» (para além de uma «frasezita do ano», que corresponde ao título da canção *O, wie schön ist Panama*: Que bonito é o Panamá!).

É geralmente reconhecido que a pós-verdade do dicionário Oxford é sobretudo encarnada por Donald Trump, pois foi o seu porta-voz que confirmou à imprensa a existência de «factos alternativos». O triunfo da equipa Trump tem mais mérito tratando-se dos Estados Unidos, primeira potência mundial em lexicografia criativa institucional. Ver para esse efeito os bem-conseguidos sintagmas da CIA («*collateral damage*» para «matança de civis» ou «*waterboarding*» para «simulação de afogamento», para não falar das «armas de destruição maciça»). Outros vetores de criatividade linguística são as consultoras das empresas e as elites oficiais, que designam, por exemplo, como «*tax ruling*» o que não é mais que um cambalacho ministerial, ou as administrações públicas que têm alardes tão poéticos como «auxiliar de ação educativa»(15), em Portugal, «*agente de la movilidad*»(16), em Espanha, ou «*vigile del fuoco*»(17), em Itália. Claro que também há lexicógrafos particulares. Na mesma linha do conseguidíssimo «*finiquito en diferido*»(18) (rescisão em diferido) de María Dolores de Cospedal (secretária-geral do PP espanhol), em 2016 houve em Espanha contributos geniais, embora a Fundéu já não tenha ido a tempo de os aproveitar. A «*contabilidad extracontable*» (contabilidade extracontabilística) ou as «*fundaciones visillo*» (fundações fachada) de Luis Bárcenas (PP espanhol) marcaram sem dúvida o «apequerruchamento» lexicográfico (dois termos de quinhentos por um de mil). Embora seja verdade que, em conjunto, ninguém como Trump apelou «à emoção e às convicções pessoais»(19) para corresponder à definição de «pós-verdade» formulada no Oxford. O trumpismo linguístico irrompeu entre as palavras candidatas aos prémios de 2016 em diferentes línguas, como o sueco: «*trumpifiering*», «*trumpeffekt*»... E o jornal *Dagens Nyheter* agradeceu a Trump a sua contribuição para a linguística. É Trump que pode corrigir o apresentador Wyoming do canal televisivo *la Sexta*: «Já conhecem a verdade; agora vamos contar-lhes os factos».

ajmleyla@gmail.com

---

(1) *puntoycoma*, http://ec.europa.eu/translation/spanish/magazine/es_magazine_en.htm.
(2) Palavra do Ano — Portugal, http://www.palavradoano.pt/.
(3) Por curiosidade, as palavras do ano de 2016 no Brasil, em Angola e Moçambique foram, respetivamente, «indignação», «crise» e «paz».
Globo: G1, *Educação*, «'Indignação' é a palavra do ano para os brasileiros, diz pesquisa»,
http://g1.globo.com/educacao/noticia/indignacao-e-a-palavra-do-ano-para-os-brasileiros-diz-pesquisa.ghtml.
Palavra do Ano — Angola, http://www.palavradoano.co.ao/#arquivo.
Palavra do Ano — Moçambique, http://www.palavradoano.co.mz/#arquivo.
(4) *man + explain:* «mansplaining — (informal, pejorative) The act of condescendingly explaining something, particularly by a man to a female listener in order to appear knowledgeable or from a mistaken presumption that she has an inferior understanding of the topic.», Wikipedia, *mansplaining*, https://en.wiktionary.org/wiki/mansplaining.
(5) Le Soir, *Débats*, «Le nouveau mot de l'année est... Brexit»,
http://www.lesoir.be/1403012/article/debats/2016-12-28/nouveau-mot-l-annee-est-brexit.
(6) *pute + à + clic:* «Se dit putaclic tout contenu ne respectant pas dans ses grandes lignes la promesse éditoriale suggérée par son accroche aguicheuse. — (Parlons Youtube — Pierre Croce)», Wikipedia, *putaclic*, https://fr.wiktionary.org/wiki/putaclic.
A palavra vencedora do *Le Soir* foi Brexit.
(7) Van Dale, *Woord van het Jaar: 2016*, http://woordvanhetjaar.vandale.be/nl/historie.html.
Samson + seks: Samson en Gert é uma série infantil de sucesso na Bélgica,
https://fr.wikipedia.org/wiki/Samson_en_Gert.

Por curiosidade, a palavra do ano de 2016 nos Países Baixos foi «*treitervlogger*» (algo como «vlóguer assediador), Van Dale, *Woord van het Jaar – Historie*, «Woord van het Jaar 2016», http://woordvanhetjaar.vandale.nl/nl/historie.html#year2016.
(8) Para «*phubbing*», termo que apenas resulta em inglês. *Phone* + *snub:* «phubbing — The act of ignoring a conversation going on around oneself to focus on one's phone.», Wikipedia, *phubbing*, https://en.wiktionary.org/wiki/phubbing.
(9) Le Festival du mot, *Mot de l'année 2016*, http://www.festivaldumot.fr/article/le-mot-de-l-annee-4fe0726ebe733.
(10) *Web* + *ebete* (idiota).
Repubblica, *Arte e Cultura*, «La parola del 2016 è stata 'webete': Mentana fa tendenza, ma spunta un'attestazione anni '90», http://www.repubblica.it/cultura/2017/01/08/news/la_parola_del_2016_e_stata_webete_mentana_fa_tendenza_anche_se_spunta_un_attestazione_anni_90-155620046/.
(11) Real Academia Galega, *Actualidade: Novas*, «'Irmandade', Palabra do Ano 2016», http://academia.gal/inicio/-/asset_publisher/m2gF/content/'irmandade'-palabra-do-ano-2016.
(12) Real Academia Galega, *Dicionario da Real Academia Galega*, «alburgada», http://academia.gal/dicionario/-/termo/busca/alburgada.
(13) Oxford Dictionaries, *Word of the Year*, «Word of the Year 2016 is...», https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016.
(14) Der Spiegel, *Kultur*, «"Postfaktisch" ist Wort des Jahres 2016», http://www.spiegel.de/kultur/gesellschaft/wort-des-jahres-2016-postfaktisch-gekuert-a-1125124.html.
(15) Contínuo (escola).
(16) Polícia de trânsito.
(17) Bombeiro.
(18) YouTube, *Cospedal intenta explicar en una comparecencia surrealista el "finiquito en diferido" de Bárcenas*, carauarta5, https://www.youtube.com/watch?v=LekyXrHFl-U.
(19) Oxford Dictionaries, *Word of the Year*, «Word of the Year 2016 is...», https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016.

A polémica escusada

Jorge Madeira Mendes
Direção-Geral da Tradução — Comissão Europeia

Periodicamente, o tema ressuscita.

Mais de quatro anos depois de adotada pelo *Diário da República* e pelos serviços de tradução portuguesa das instituições europeias; utilizada correntemente pelas três maiores estações de televisão e pela maior parte da imprensa escrita de Portugal; ensinada desde há anos nas escolas (o que significa haver já uma geração de jovens portugueses que não sabem escrever de outra maneira) — a ortografia resultante do Acordo Ortográfico de 1990 (AO90) continua a ser cavalo de batalha numa polémica que tem tanto de estéril quanto de desgastante.

Mesmo depois de demonstrado o caráter falacioso da argumentação que Vasco Graça Moura inaugurou e outros aproveitaram — um «facto» não é um «fato», um «pacto» não é um «pato», um «cágado» não se confunde com algo malcheiroso —, figuras públicas da sociedade portuguesa (da parte das quais seriam de esperar maior rigor e conhecimentos mais exatos) continuam a repescá-la.

José Pacheco Pereira — que emprega palavras como *know-how*, *bullying*, *media*, *likes*; que recorre a termos como «instrumental» (colagem ao inglês *instrumental*, que, em português, se traduziria por «útil»), «municipalidade» (colagem ao francês *municipalité*, que, em português, se traduziria por «município»), «adquirido» (colagem ao francês *acquis*, que, em português, se traduziria por «acervo» ou por «conquistas»); que utiliza o verbo «falar» em lugar do verbo «dizer»; que transforma verbos reflexos em verbos intransitivos («a situação começa a acalmar»); que emprega neologismos escusados como «exce[p]cionar» (descurando o consagrado «exce[p]tuar»); que constrói frases à inglesa («Várias prisões foram feitas... notícias contraditórias chegavam à polícia», quando, classicamente, o português usaria inversões: «Foram feitas várias prisões... chegavam à polícia

notícias contraditórias»)[1]; e que, num programa da série *Quadratura do Círculo*, ouvi explicar que «adolescentes» queria dizer *teenagers* (não, não foi o contrário!) — compara o AO90 a uma «perda de identidade» e a um «abastardamento da língua». Não discuto os méritos de Pacheco Pereira enquanto historiador, mas discuto as suas incongruências quando declara um quixotesco combate ao que não passa de mero ajuste da ortografia (como tantos que houve ao longo da história da língua e que consiste, maioritariamente, em se suprimirem consoantes mudas)[2]. Causas mais gloriosas não lhe faltariam.

Considerar que um ajuste da ortografia, baseado essencialmente na supressão de consoantes mudas, constitui um «abastardamento», uma «perda de identidade», é no mínimo deplorável — sobretudo se atendermos a que os próprios falantes da língua portuguesa estão a transformá-la numa espécie de esperanto atabalhoado[3], invadindo-a insolentemente com palavras americanas para as quais o nosso léxico contém alternativas consagradas (logo, mais inteligíveis). Amiúde por puro exibicionismo, diz-se *target* em vez de «alvo» ou «meta», *coach* em vez de «treinador», *training* em vez de «formação», *break* em vez de «intervalo», *benchmarking* em vez de «aferição comparativa», *assets* em vez de «ativos», *players* em vez de «intervenientes», *chef* em vez de «chefe», *shopping center* em vez de «centro comercial», *sponsoring* em vez de «patrocínio», *deadline* em vez de «prazo», *stakeholder* em vez de «parte interessada», *turning point* em vez de «ponto de viragem», *meeting point* em vez de «ponto de encontro», *submeter* em vez de «enviar», «entregar» ou «apresentar» (por tradução acrítica do inglês *to submit*), *suportar* em vez de «apoiar» (por tradução acrítica do inglês *to support*)... e fico-me por apenas alguns exemplos do verdadeiro «abastardamento», da verdadeira «perda de identidade».

Do mesmo modo, se seria descabido contestar os méritos de António Vitorino de Almeida enquanto compositor, permito-me pôr em causa uma afirmação sua, num programa da série *5 para a Meia-Noite*, há um par de anos, de que a palavra «espetador» se confundia com «esp'tador»: lembro que «esquecer», «arrefecer» e «padaria» não se pronunciam «esq'cer», «arref'cer» ou «pɐdaria» — ou seja, há vogais que se «abrem», independentemente de a sua oralidade ser assinalada por acentos ou outros diacríticos. Se para abrirmos o segundo «e» de «espetador» tivermos de lhe acrescentar um «c» mudo, então teria lógica escrevermos «esque**c**cer», «arrefe**c**cer» e «pa**c**daria», subentendendo-se que aqueles «cc» seriam mudos e só serviriam para abrir a respetiva vogal antecedente.

Por sua vez, José Pedro Vasconcelos (apresentador do dito *5 para a Meia-Noite*) ironizou com a (por ele suposta) nova ortografia «egícios» (brilhante suposição a que terá chegado na sequência de o AO90 ter introduzido a grafia «Egito»). Ora, nunca o AO90 impôs a aberração «egícios», porquanto o seu objetivo não é alterar a pronúncia, mas sim adaptar a escrita à pronúncia — logo, se na pronúncia de «egípcios» há um «p», também na sua ortografia continua a havê-lo.

Ricardo Araújo Pereira, ao mesmo tempo que apela a que «não toquem na [sua] língua», sustenta que a grafia «arquitetas» implica, por força, o fechamento da vogal «e» para [*ê*], com a inevitável (e jocosa) ambiguidade de sentido que daí resultaria. Ora, o fechamento do «e» em «arquitetas» é tão forçoso como em «netas» (a ninguém passa pela cabeça que, na frase «a avó foi passear as netas», a última palavra se pronuncie *nêtas*). Em português, independentemente do AO90, verificam-se há muito ambiguidades ortográficas, em consequência das quais nem sempre é evidente o fechamento ou a abertura de vogais: por exemplo, na palavra «segredo», o segundo «e» será aberto [*è*] se se tratar de uma das formas do verbo *segredar*, mas será fechado [*ê*] se se tratar de um substantivo («eu segredo um segredo»); na palavra «almoço», o segundo «o» será aberto [*ò*] se se tratar de uma das formas do verbo *almoçar*, mas será fechado [*ô*] se se tratar de um substantivo («eu almoço à hora de almoço»); na palavra «cor», o «o» será aberto [*ò*] se se tratar do substantivo designativo de memória (por «coração»), mas será fechado [*ô*] se se tratar do substantivo designativo da impressão cromática («conheço de cor aquela cor»); na palavra «topo», o primeiro «o» será aberto [*ò*] se se tratar de uma das formas do verbo *topar*, mas será fechado [*ô*] se se tratar de um substantivo («eu topo tudo no topo»); na palavra «pega», o «e» será aberto [*è*] se se tratar de uma das formas do verbo *pegar*, mas será fechado [*ê*] se se tratar do substantivo designativo de certos corvídeos («pega aí a pega»); o mesmo com a palavra «exagero», na qual o segundo «e» será aberto [*è*] se se tratar de uma das formas

do verbo *exagerar*, mas será fechado [*ê*] se se tratar de um substantivo («eu exagero com grande exagero»); o mesmo com a palavra «coma», na qual o «o» será fechado [*ô*] se se tratar de uma das formas do verbo *comer*, mas será nasalado [*õ*] se se tratar do substantivo designativo de um estado de inconsciência («ora coma, se não quer entrar em coma»). E muitos mais exemplos haveria. Só pelo sentido e pelo contexto podemos decidir-nos por uma pronúncia ou pela outra.

No fundo, a oposição às novas regras ortográficas do português assenta, maioritariamente, em equívocos e fantasmas. A ortografia do português sofreu constantes reformas ao longo da História. A quem passa pela cabeça que, no tempo de D. Dinis, se escrevia como imediatamente antes deste chamado «Acordo Ortográfico de 1990»? E nem precisaríamos de recuar à época de D. Dinis, bastando quedarmo-nos pela de Camilo Castelo Branco (ou melhor, «Camillo Castello-Branco»).

As regras pelas quais se pautam os mais acérrimos adversários do AO90 são as de uma reforma de 1945 (revista em 1973). E já antes dessa houvera a de 1911, causadora de violentíssimas polémicas nos jornais da época, com desafios para duelo e ameaças de bengalada. Nesse tempo, escrevia-se «lyrio», «mechanico», «photographico», «auctor», «Christo», «portugueza», «illustração». Fernando Pessoa ateve-se sempre à «pharmacia» e ao «cagado», porque nessa altura as esdrúxulas não levavam acento gráfico (ou «graphico»?).

Já ouvi, em inquérito de rua, uma telespetadora dizer, em categórico tom proclamatório, que estava a ler o último livro de Saramago em inglês, porque não suportava lê-lo «em acordês». Seria caso para lhe perguntar se também lia Pessoa numa qualquer tradução: é que, no seu tempo, Pessoa se ergueu acerrimamente contra a reforma ortográfica de 1911, por via da qual a escrita do português sofreria uma transformação muito mais drástica do que a atual. Acresce que, antes de chegar à versão imediatamente anterior à do AO90, a ortografia sofreu ainda a reforma de 1945. Portanto, o Pessoa que hoje a dita telespetadora eventualmente lê distancia-se muito mais do Pessoa original do que o Saramago pré-AO90 se distancia do Saramago pós-AO90.

Um outro argumento, dos mais conspícuos, tem a ver com a grafia, imposta pelo AO90, da terceira pessoa do singular do verbo «parar» no indicativo presente: «ele/ela pára» passou para «ele/ela para». Portanto, na frase, exemplificativa, «chuva para Lisboa», ficar-se-ia sem saber se se prevê chuva para a cidade de Lisboa ou se esta última ficou parada devido a uma copiosa precipitação. É certo, mas eu atrever-me-ia a alvitrar que a ambiguidade da frase «chuva para Lisboa» poderá ser vista mais como um trocadilho apetecível do que como uma inconveniência a evitar. Quero dizer, concretamente, que um título de jornal como aquele poderia ser propositado, precisamente para incitar à leitura da notícia. Seria, também, o caso de um título como «Empresários com sede no deserto»: o leitor sentir-se-ia tentado a procurar no texto da notícia se se tratou de empresários que, em viagem pelo deserto, não se tinham devidamente munido de líquidos, ou de empresários que sediaram as suas empresas (de safáris, por exemplo) no próprio deserto.

Ainda no âmbito das ambiguidades inócuas: alguns dos leitores talvez se lembrem de uma peça de teatro, de há umas boas dezenas de anos, intitulada *Um zero à esquerda*; havia ali uma ambiguidade propositada, porquanto esta expressão tanto pode ser tomada, seriamente, no seu significado matemático como, maliciosamente, com um sentido político.

A ambiguidade entre «para» [*à*] e «para» [*â*] não é mais grave do que a ambiguidade — há muito existente — entre «sede» [*ê*] e «sede» [*è*]. A diferença é que a possível ambiguidade entre «para» [*à*] e «para» [*â*] é uma novidade resultante do AO90, ao passo que a possível ambiguidade entre «sede» [*ê*] e «sede» [*è*] já existia (tal como muitas outras), por imposição do Acordo Ortográfico de 1945, quando aprendemos a escrever.

E note-se que o «problema» não se restringe à língua portuguesa. Um dia, deparei-me com um texto relativo ao livro *The cosmic race* (versão inglesa de uma obra do mexicano José Vasconcelos Calderón). Tratar-se-ia de uma «*corrida* cósmica»? Só ao cabo de alguns parágrafos me dei conta de que o autor se referia aos povos iberoamericanos, de língua espanhola e portuguesa, que designava por

«*raça* cósmica», porque, em sua opinião, resultavam de um cadinho de todas as raças e etnias que compõem a Humanidade.

Rematando a minha refutação de que a ambiguidade resultante de uma homografia para as palavras «para» [forma verbal] e «para» [preposição] constitua um problema intransponível, citarei um exemplo extraído de um jornal ecologista poucos anos após a democratização de 25 de abril de 1974: em determinado artigo escolheu-se o título «*Uma experiência de recuperação de bufos*»; à primeira vista, houve quem se interrogasse se a experiência tinha a ver com a «reciclagem» dos denunciantes a que recorria a recentemente extinta polícia política; ora, na verdade, tratava-se de recuperar (para o *habitat* natural) espécimes, por algum motivo a ele subtraídos, da espécie *Bubo bubo*, que os especialistas designam como «mocho-real» mas que, em português vernáculo, tem o nome genérico de «bufo». A ambiguidade fora propositada, no sentido de «apimentar» o título e, desse modo, despertar a curiosidade para o texto.

Retomo agora o «Egito» e os «egípcios», para discutir um outro frequente argumento dos opositores à nova ortografia: todas as palavras de uma mesma família etimológica deveriam conservar as consoantes de origem, ainda que não se pronunciassem. Assim, o topónimo «Egipto» deveria continuar a escrever-se com «p» (embora não pronunciado) por uma espécie de solidariedade etimológica com o correspondente gentílico, «egípcio». Ora, isto é esquecer que, só no domínio dos topónimos e respetivos gentílicos, há inúmeros casos de divergência, com o gentílico a manter-se mais próximo do étimo (latino ou latinizado, na maioria dos casos). Com efeito, dizemos «Braga» mas «bracarense», «Bragança» mas «brigantino», «Chaves» mas «flaviense», «Évora» mas «eborense», «Guimarães» mas «vimaranense», «Idanha» mas «egitanense», «Lamego» mas «lamecense», «Lisboa» mas «lisbonense», «Penafiel» mas «penafidelense».

É que, em tempos, havia uma cidade chamada «Bracara», cujos habitantes eram, logicamente, «bracarenses»; entretanto, o topónimo evoluiu para «Braga» mas o gentílico continuou a ser «**bracar**ense». Do mesmo modo, em tempos, havia uma cidade chamada «Brigantia», cujos habitantes eram, logicamente, «brigantinos»; entretanto, o topónimo evoluiu para «Bragança» mas o gentílico continuou a ser «**brigant**ino». Havia uma cidade chamada «(Aquae) Flaviae», cujos habitantes eram, logicamente, «flavienses»; entretanto, o topónimo evoluiu para «Chaves» mas o gentílico continuou a ser «**flavi**ense». Havia uma cidade chamada «Ebora», cujos habitantes eram, logicamente, «eborenses»; entretanto, o topónimo evoluiu para «Évora» mas o gentílico continuou a ser «**ebor**ense». Havia uma localidade chamada «Vimaranes» (propriedade da condessa Mumadona Dias), cujos habitantes eram, logicamente, «vimaranenses»; entretanto, o topónimo evoluiu para «Guimarães» mas o gentílico continuou a ser «**vimaran**ense». Havia uma cidade chamada «Egitania», cujos habitantes eram, logicamente, «egitanenses»; entretanto, o topónimo evoluiu para «Idanha» mas o gentílico continuou a ser «**egitan**ense». Havia uma cidade chamada «Lamecum», cujos habitantes eram, logicamente, «lamecenses»; entretanto, o topónimo evoluiu para «Lamego» mas o gentílico continuou a ser «**lamec**ense». Havia uma cidade chamada «Olisippo», que evoluiu sucessivamente para «Olissipona» e «Lisbona»; os habitantes de «Lisbona» eram, logicamente, «lisbonenses»; entretanto, o topónimo evoluiu para «Lisboa» mas o gentílico continuou a ser «**lisbon**ense» (que coexiste com a forma mais popular «lisboeta»). E havia uma localidade chamada «Pena Fidelis», cujos habitantes eram, logicamente, «penafidelenses»; entretanto, o topónimo evoluiu para «Penafiel» mas o gentílico continuou a ser «**penafidel**ense».

Pela mesma ordem de ideias, seriam «ilógicas» as formas «Chipre»/«cipriota». O topónimo original, «Kýpros», teve uma evolução distinta da do correspondente gentílico. É também o caso de «Salamanca»/«salmantino».

No caso em apreço, temos um país, de capital no Cairo, geograficamente coincidente com um império da Antiguidade chamado «Egipto», cujos habitantes eram, logicamente, «egípcios». Entretanto, o topónimo evoluiu para «Egito» mas o gentílico continuou a ser «**egíp**cio». Se há aqui contrassenso, então rebatizem-se imediatamente Braga, Bragança, Chaves, Évora, Guimarães, Idanha, Lamego, Lisboa e Penafiel com os seus nomes de há mil ou dois mil anos. Insistir em escrever «Egipto» embora

pronunciando «Egito» faz tanto sentido como insistir em escrever «Bracara» embora pronunciando «Braga».

O artigo *Divergências na família*, que a seguir publico, enuncia uma série de estirpes etimológicas, distintas dos topónimos e gentílicos, em cada uma das quais é notória a deriva das palavras mais usadas (e, portanto, mais «desgastáveis») relativamente às palavras-irmãs: embora provindo de um étimo comum (regra geral, latino ou latinizado), as palavras mais utilizadas são também mais erodidas (evolução por via popular), num fenómeno que se traduz pelo abrandamento de consoantes («p» para «b», «t» para «d», «c[k]» para «g», «c[s]» para «z»), pela palatalização de pares de consoantes («cl»/«fl»/«pl» para «ch»/«lh»), pela aglutinação de sílabas, enfim, pela supressão pura e simples de consoantes, ao passo que as palavras menos utilizadas da mesma linhagem se mantêm muito mais próximas do étimo, com a adaptação idiomática que a evolução por via erudita constitui.

Jorge-Madeira.Mendes@ec.europa.eu

---

(1) Pereira, J. P., *Álvaro Cunhal, Uma Biografia Política: O Secretário-Geral*, vol. 4, Temas e Debates, 2015, ISBN 9789727598052.
(2) O AO90 altera 1,6% das palavras do português de Portugal e 0,5% das do português do Brasil, Pereira, A, *Acordo Ortográfico: O que mudou*, «.dicção, dição ou tanto faz?», https://acordo-ortografico.blogspot.be/2012/10/?m=0.
(3) Sem desprimor para o inventor do esperanto, que Zamenhof não poderia adivinhar certas tendências iníquas numa língua do oeste ibérico pouco mais de um século depois.

Divergências na família

Jorge Madeira Mendes
Direção-Geral da Tradução — Comissão Europeia

Este artigo, complementar ao precedente («A polémica escusada»), apresenta grupos de palavras etimologicamente aparentadas que revelam discrepância evolutiva: as palavras mais utilizadas — logo, mais sujeitas a erosão — afastam-se visivelmente do étimo (latino, em geral); as outras mantêm-se mais próximas.

abelha, mas **api**cultor;
abismo, mas **abiss**al;
agudo, mas **acut**ilante;
agulha, mas **acíc**ula;
alma, mas **ânim**o, **anim**ado;
água, mas **aqu**oso, **aqu**ífero;
areia, mas **aren**oso, **aren**ito, **aren**a;
ausente, ausência, mas **absent**ismo;
bem, mas **bene**ficência, **bene**mérito;
bexiga, mas **vesícula**;
bispo, mas **episcop**al;
bom, boa, mas **bon**ificação, **bon**omia, a**bon**amento;
cadeira, mas **cátedra**, **catedr**al;
cair, mas (de)**cad**ência, (de)**cad**ente, **cad**uco, **cad**ucidade;
cativar, cativo, mas **capt**ar, **capt**ação;
ceia, mas **cená**culo;
céu, mas **cel**este, **cel**estial;
chama, mas **flam**ejante, in**flam**ar, in**flam**atório, in**flam**ação;
chamar, mas **clam**ar, pro**clam**ação, re**clam**ar, re**clame**;
chão, chã, lhano, lhaneza, lhanura, mas **pl**ano, **pl**anura;
chave, mas **cl**ave;

cheio, mas **plen**o, **plen**itude;
choro, chorar, mas **pl**angente, im**plor**ar;
chumbo, mas **plúmb**eo;
chuva, chuvoso, chover, mas **pluvi**osidade, **pluvi**oso, **pluvi**ómetro;
coalho, coalhar, mas **coágul**o, **coagul**ar;
cobra, mas **colubr**ear, **colubr**ídeo, **colubr**ino;
cobre, mas **cuprí**fero;
colheita, mas **cole(c)t**a, **cole(c)t**ânea;
conceito, conceição, mas **conce(p)**ção, **concep**tual;
conhecer, conhecimento, mas **cogn**itivo, **cogn**oscível, (in)**cógn**ito;
cor, mas **color**ido, **color**ação;
crer, crente, incréu, mas **cred**o, **créd**ito, **cred**ível, (in)**créd**ulo;
cruel, crueldade, mas **crudel**íssimo, **crudel**izar;
dedo, mas **dígit**o, **digit**al;
dicionário, mas **dicç**ão;
dizer, mas **dicç**ão;
doce, doçura, mas **Dulce**, **dulc**íssimo, e**dulc**orante;
efeito, mas **efe(c)t**ivo;
emprenhar, prenhe, mas im**pregn**ar;
entre, mas **inter**ativo, **inter**nacional;
estômago, mas **estomac**al;
fado, mas **fat**al, **fat**alidade, **fat**alismo;
fala, falar, falante, mas **fábula**, **fabula**r, **fabula**ção, **fabul**oso;
fêmea, mas **femin**ino, **femin**il, e**femin**idade;
feito, mas **fact**o, **fact**ual;
fiel, mas **fidel**idade, **fidel**íssimo, **fidel**izar;
filho, filha, mas **fili**al;
fogo, mas **foc**o, **foc**al, **foc**ar;
fresta, mas de**fenestr**ar, de**fenestr**ação;
frio, mas **frígid**o, **frigid**íssimo;
geral, mas **gener**alidade, **gen**érico;
gerar, geratriz, mas **gener**ativo, **gener**atriz, re**gener**ar;
hoje, mas **hod**ierno;
idade, mas **et**ário, co**et**âneo;
igreja, mas **eclesiá**stico;
igual, igualdade, igualitário, mas **eq**uidade, **eq**uação, **eq**uitativo;
ilha, mas **insul**ar, pen**ínsula**, pen**insul**ar;
inteiro, inteirar, inteireza, mas **íntegr**o, **integ**ridade, **integ**rar, **integ**rismo;
jazer, jazigo, jazida, mas **jac**ente;
jogo, jogar, mas **joc**oso;
lago, lagoa, mas **lagun**a, **lagun**ar, **lacun**a, **lacu**stre;
lágrima, mas **lac**rimal, **lac**rimejar, **lac**rimejante;
leite, mas **lác**teo;
lenda, lendário, mas **legenda**, **legend**ário;
letra, mas **liter**al, **liter**ário, **liter**acia;
livre, mas **liber**dade, **libér**rimo, **liber**al, **liber**tário;
louvor, louvar, louvável, mas **laud**atório, **laud**ativo, **laud**ável;
macho, mas **másc**ulo, **masc**ulino, e**masc**ular;
madeira, mas **matéria**;
mãe, mas **mater**no, **mater**nal, **mater**nidade;
mês, mas **mens**al, **mens**trual;
milagre, mas **miracul**oso;
moeda, mas **monet**ário;
nau, náutica, mas **nav**e;
neto, mas **nepot**ismo;
nuvem, mas **nebul**oso, **nebul**osidade;
oito, oitavo, (O)outubro, mas **oct**ogésimo, **oct**ogenário, **oct**ógono, **oct**ópode, **oct**ossílabo, **oct**ostilo;
olho, mas **ócul**o, **ocul**ar;
orelha, mas **aurícula**, **auricul**ar;
ouro, mas **aur**ífero;

ouvir, mas **audí**vel, **audi**tivo, **audi**tório;
pai, mas **pater**no, **pater**nal, **pater**nidade, **pátr**ia, **pátr**io, **patr**iota;
parede, mas **pariet**al;
pé, mas **ped**al, **ped**estal, **ped**estre, **pedí**culo, **ped**icura, **ped**únculo;
pedir, mas **peti**ção;
pedra, mas **pétr**eo, **petr**óleo, **petr**ificar;
peixe, mas **pisc**ícola, **pisc**icultura, **pisc**atório, **pisc**iforme;
perigo, perigoso, mas **pericul**osidade;
peso, pesar, mas **pond**eral, **pond**erar;
pessoa, pessoal, mas **person**alidade, **person**agem, **person**ificar;
pó, mas **pulver**ulento;
poder, mas **pot**ência, **pot**ente, **pot**entado, **pot**estade;
povo, mas **pop**ular;
prova, comprovativo, mas **prob**atório;
raiz, mas **radic**ar, **radic**al, er**radic**ar;
rei, reino, reinado, mas **reg**er, **reg**ente, inter**regn**o, **rég**ulo, **rég**io;
respeito, mas **respe(c)t**ivo;
rim, mas **ren**al;
sabão, mas **saponi**ficação;
saber, sábio, sabedoria, mas **sap**iência, **sap**ientíssimo;
sabor, mas sá**p**ido, insí**p**ido;
sagrado, sagrante, mas **sacr**atíssimo, **sacr**amento, **sacr**ário, **sacr**a, **sacr**ifício, **sacr**ilégio, **sacr**istão;
sair, mas **sali**ente, **sali**entar, **sali**ência;
saúde, saudável, mas **salut**ar;
seio (paranasal), mas **sinus**ite;
seita, mas **sect**ário, **sect**arismo, **sec**ção, **sec**cionar, **sec**cionamento;
sete, mas **sept**uagenário;
só, mas **sol**idão;
tábua, tabuada, mas **tabul**ar, **tabul**ário, **tabul**eta, **tabul**ista;
teia, mas **tela**;
ter, mas **ten**ente;
touro, mas **tau**romaquia;
tráfego, mas **tráfic**o;
trevas, mas **tenebr**oso;
veia, mas **ven**oso, **ven**al;
véu, mas **vel**ado, **vel**ar, **vel**o (?);
víbora, mas **vip**erino, **vip**erídeo;
vida, mas **vit**al, **vit**alício, **vit**alidade...

...e um grande «etc.», que inclui, designadamente, todos os adjetivos/substantivos terminados em «-ável» ou «-ível» («amável», «possível») e aos quais corresponde um substantivo, designativo de estado, terminado em «-abilidade» ou «-ibilidade» («ama**bil**idade», «possi**bil**idade»).

Jorge-Madeira.Mendes@ec.europa.eu

Por uma homografia decente

Maria Manuel Monteiro Ricardo
Técnica superior/tradutora jurídica — Direção-Geral da Política de Justiça/Ministério da Justiça

O Acordo Ortográfico (AO) de 1945, ratificado por Portugal, mas rejeitado pelo Brasil, era genericamente de tipo etimológico. Vultos importantes da cultura portuguesa interessavam-se pelo tema da ortografia. Fernando Pessoa/Bernardo Soares, no *Livro do Desassossego*, foi uma voz empenhada nessa causa na década trinta do século XX, como mostra o texto de que cito um excerto:

> Não tenho sentimento nenhum politico ou social. Tenho, porém, num sentido, um alto sentimento patriotico. Minha patria é a lingua portuguesa. Nada me pesaria que invadissem ou tomassem Portugal, desde que não me incommodassem pessoalmente. Mas odeio, com odio verdadeiro, com o unico odio que sinto, não quem escreve mal portuguez, não quem não sabe syntaxe, não quem escreve em orthographia simplificada, mas a pagina mal escripta, como pessoa própria, a syntaxe errada, como gente em que se bata, a orthographia sem ípsilon, como escarro directo que me enoja independentemente de quem o cuspisse.
> Sim, porque a orthographia também é gente. A palavra é completa vista e ouvida. E a gala da transliteração greco-romana veste-m'a do seu vero manto régio, pelo qual é senhora e rainha.(1)

O AO de 1990, ratificado por Portugal, Cabo Verde, São Tomé e Príncipe e Brasil (só Angola e Moçambique não o tinham feito ainda), quebrou este percurso, com omissão das consoantes mudas, o oposto à citada «transliteração greco-romana» de Fernando Pessoa e privilegiou a grafia de tipo tendencialmente fonético, expressa numa síntese perfeita: «se se pronuncia, fica, se não se pronuncia, cai». Este foi um lado positivo do AO.

O outro lado positivo foi a conceção de que nós não comunicamos por palavras isoladas, comunicamos por frases, da frase se passa ao contexto e é o contexto que nos dá o significado:

> — Uma **colher** de açúcar?
> — Duas, por favor, **gosto** do café açucarado.
> — Eu **gosto** do café a saber a café, não **gosto** do café a saber a açúcar, muito menos do açúcar a saber a café.
> — O **gosto** de cada um não se discute.
> — Chegou atrasado, apesar de ter agora um belo carro.
> — Sou um homem de negócios, estou a **colher** os frutos do trabalho e dos investimentos na fábrica de Palmela.
> — Tem **olho** para o negócio.
> — Sim, **olho** pela minha fábrica, é ela que me enriquece.
> — **Para** muito tempo na fábrica?
> — Estou lá das 9 às 19h, há sempre coisas **para** tratar, assuntos a resolver.

As palavras em negrito são homógrafas porque, tendo a mesma grafia, são diferentes na categoria morfológica (verbo e nome), são «mui decentes» (embora «para muito na fábrica …» não o seja tanto), no sentido etimológico, patente n'*Os Lusíadas* (mui decente = mui conveniente), como se vê no texto que passo a citar:

> Formosa filha minha, não temais
> Perigo algum nos vossos Lusitanos (…)
> Que eu vos prometo, filha, que (…)
> Vereis a terra que a água lhe tolhia,
> Que inda há de ser um porto mui **decente**,
> Em que vão descansar da longa via
> As naus que navegarem do Ocidente.(2)

A frase atrás citada «**Para** muito tempo na fábrica?», aparentemente inocente, levanta um problema relevante.

Por um lado, para além desta frase, há um sem número de outras em que é fácil ler **para** (v) de **para** (p):

> «Para fazeres, em segurança, a viagem, para ali.»
> «Se receberes um SMS no teu telemóvel, não atendas, encosta, para e ouve.»
> «És alcoólico compulsivo?, para de acusar os teus amigos, acusa-te a ti próprio.»
> «Estás cansada, por isso, para de trabalhar, hoje.»
> «O elétrico para para as pessoas descerem e subirem.»
> «Porque é que para as pessoas passarem, para o trânsito?»

> «Temos muita informação, informação excessiva que, às vezes, para para informar que o FC Bayern München ganhou ou o Cristiano Ronaldo do Real Madrid marcou.»

Por outro lado, há casos, embora excecionais, que pertencem à língua que lemos, em que o leitor não sabe se deve ler «para» como verbo ou como preposição, exemplos:

> Para o trabalho que tinhas começado e para o outro que o mestre indicou, pois há vida além do trabalho, paráfrase de uma frase célebre «há mais vida para além do défice» de um célebre político, embora custe dizê-lo, pois pode revoltar o trabalhador, tantas vezes privado do seu direito ao descanso, é preciso mais empenho.

Só quando lemos «é preciso mais empenho», ficamos a saber que **para** é preposição. Seria diferente, se a frase fosse:

> Para o trabalho que tinhas começado e para o outro que o mestre indicou, pois há vida para além do trabalho, paráfrase de frase célebre de um célebre político «há vida para além do défice», embora custe dizê-lo, pois pode prejudicar a empresa, tantas vezes prejudicada por greves e baixas por motivo de saúde, está na hora de parar.

Só quando lemos «está na hora de parar», é que ficamos a saber que **para** é verbo.

Poder-se-á argumentar que os exemplos apresentados são esteticamente insustentáveis, mas replicar-se-á que o que nos move, aqui e agora, é um tema linguístico e não estilístico.

Haveria, em teoria, uma maneira de distinguir, e é a de, à semelhança de **pôr** / **por**, acentuar sempre o verbo (**pára**) para o distinguir da preposição (**para**). Isso não foi pensado pelos mentores do AO e, por isso, não está previsto no Acordo.

Mas o AO não é um texto sagrado. Com um acordo entre os que o adotaram, poder-se-ia acabar com a referida homografia indecente para/para, para a transformar em homografia decente **pára/para**.

Porém, se não foi possível fazê-lo em tempo certo, será possível fazê-lo, hoje, com o comboio em andamento, quando se antepõem outras agendas aparentemente mais importantes (economia, fiscalidade, educação, justiça, etc.) e mais urgentes que um problema linguístico?

Resta a «pressão» da sociedade civil (jornalistas, cronistas, escritores, editores, leitores, falantes) sobre a política (partidos, parlamentares, governo), da Sociedade da Língua Portuguesa, vocacionada para os problemas da Língua, da Academia das Ciências de Lisboa, da Comunidade Linguística, dos tradutores da Administração Pública e dos tradutores de português das instituições da União Europeia, para se empenharem, todos juntos, todos um, por uma boa causa, por uma homografia decente, escrevendo sempre **pára** (v) / **para** (p), à imagem e semelhança do já referido **pôr** (v) / **por** (p**)**.

mmanuel@dgpj.mj.pt

---

(1) Soares, B., *Livro do Desassossego*, https://pt.wikisource.org/wiki/A_minha_p%C3%A1tria_%C3%A9_a_l%C3%ADngua_portuguesa.
(2) Camões, L. de, *Os Lusíadas*, canto II, estrofes 44 e 48, https://pt.wikisource.org/wiki/Os_Lus%C3%ADadas/II.

Orto90 — Converter os textos da União Europeia para o AO90… e não só

Hilário Leal Fontes
Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

[Artigo apresentado originalmente no techLING — 1.º Congresso Internacional em Línguas, Linguística e Tecnologia, Braga, 12-14 de outubro de 2016, Universidade do Minho e Biblioteca Lúcio Craveiro da Silva.]

Para reutilizar as memórias de tradução com a ortografia reformada os serviços de tradução da União Europeia (UE) decidiram converter para a nova ortografia as quase duas dezenas de milhões de segmentos existentes em língua portuguesa nas memórias Euramis(1) no final de 2011. Da análise e do tratamento desse *corpus* resultou uma extensa lista de palavras a converter (lista de conversão) para a nova ortografia.

Ironia do destino, as memórias acabaram por nunca ser convertidas, mas a ferramenta baseada na lista de conversão foi rapidamente adotada pelo serviço MT@EC, a plataforma de tradução automática que serve toda a administração europeia e está disponível para as administrações nacionais dos Estados-Membros da UE e de alguns Estados do Espaço Económico Europeu. A lista de conversão já teve várias pequenas atualizações, mas em 2016 decidimos revisitar o *corpus* de língua portuguesa (que, entretanto, duplicou de volume) e a metodologia, com vista a uma nova atualização mais substancial e a preparar o terreno para atualizações periódicas. A ferramenta em si não sofreu mudanças e demos-lhe agora o nome de Orto90.

O contexto

> As instituições, órgãos e organismos da União Europeia decidiram aplicar, a partir de 1 de janeiro de 2012, o Acordo Ortográfico da Língua Portuguesa de 1990. A partir dessa data, os textos publicados no *Jornal Oficial da União Europeia* serão redigidos segundo as regras da nova ortografia, admitindo-se um período inicial de coexistência das duas ortografias.
>
> [Aviso publicado durante o mês de dezembro de 2011 no Jornal Oficial (séries L e C)]

As instituições, órgãos e organismos da União Europeia (UE) decidiram aplicar as regras do Acordo Ortográfico de 1990 (AO90) — variante português de Portugal — nas suas publicações em língua portuguesa a partir de **1 de janeiro de 2012**, coincidindo com a data de início da aplicação do AO90 pelo governo português e por todos os serviços, organismos e entidades na dependência do governo, bem como no *Diário da República*.

Em sintonia com a Resolução do Conselho de Ministros n.º 8/2011(2), foi decidido igualmente pelos serviços de tradução das instituições europeias adotar o ***Vocabulário Ortográfico do Português***(3) (VOP), do Instituto de Linguística Teórica e Computacional (ILTEC), como vocabulário oficial do AO90, primando sobre interpretações divergentes do AO90 constantes noutros vocabulários, dicionários e conversores ou corretores ortográficos.

Por questões de harmonização, no caso das **novas duplas grafias** previstas para o português de Portugal seguiram-se as predefinições do corretor ortográfico FLiP da Priberam. O FLiP era o corretor ortográfico instalado nos computadores dos serviços de tradução e do Serviço das Publicações da União Europeia — responsável pela publicação do *Jornal Oficial da União Europeia* — e, como pudemos apurar, nos computadores da Imprensa Nacional-Casa da Moeda — responsável pela publicação do *Diário da República*.

Em todos os casos, prevalecerão sempre as regras consagradas no nosso *Código de Redação Interinstitucional*(4), entretanto adaptado ao AO90, como é, por exemplo, o caso das palavras formadas por prefixação quando o segundo elemento é um estrangeirismo: anti-*dumping*, anti-*trust*, etc.(5)

O *Guia do Tradutor*[6] do Departamento de Língua Portuguesa (DLP) da Direção-Geral da Tradução (DGT) da Comissão Europeia trata também de aspetos muito concretos de ortografia, tal como vários artigos publicados no boletim da língua portuguesa das instituições europeias — «a folha»[7].

A base terminológica IATE tem igualmente estado a ser revista, adaptando o seu conteúdo à nova ortografia, de forma a poder ser usada também como ferramenta de referência para questões ortográficas.

Outra questão amplamente debatida foi a ortografia a utilizar ao fazerem-se **citações** de atos publicados antes de 2012. Com a ajuda do *Código Comercial Português* a questão ficou rapidamente resolvida. Se o *Codigo Commercial Portuguez* de 1888, ainda em vigor, já é há muito referido como Código Comercial Português, também a Directiva Arquitectos deve passar a ser referida como Diretiva Arquitetos, mesmo datando de 1985. Por outras palavras, o AO90 iria evidentemente aplicar-se não só aos novos textos como aos textos já publicados. Isto é, as **memórias de tradução** — um espólio constituído essencialmente pela legislação europeia em vigor e pelas traduções realizadas desde meados da década de 1990 (no caso da Comissão, desde 1994) — podiam e deviam ser adaptadas à nova ortografia.

O arranque

Para além da necessária **formação dos tradutores**, havia que adaptar também as ferramentas informáticas de trabalho.

Para arrancar com a nova ortografia, a DGT já tinha o **corretor ortográfico** FLiP. Bastou ativar o AO90, mantendo os valores por omissão para as novas duplas grafias, correspondentes às predefinições do FLiP. Procurou-se assim que os redatores em língua portuguesa (os tradutores da DGT) utilizassem a mesma grafia para essas palavras, independentemente da pronúncia de cada um.

Mas para não estar a atualizar constantemente a grafia de frases recorrentes traduzidas antes da aplicação do AO90, era preciso converter também as **memórias de tradução**. Para o efeito, fizeram-se primeiro testes com os dois **conversores ortográficos** disponíveis no ambiente informático da DGT: Lince e Priberam. Rapidamente nos demos conta de que as conversões oferecidas pelos dois conversores não nos satisfaziam totalmente, essencialmente por haver muitos fenómenos de subconversão (p. ex. erros, gralhas ou palavras que não respeitavam o AO45) e de sobreconversão (p. ex.: palavras estrangeiras com grafias coincidentes com o AO45) das palavras das memórias.

Partimos então para uma outra solução em que pudéssemos controlar as conversões. Depois de várias experiências decidimos que a melhor forma de atacar o problema era uma **lista de palavras a converter** (lista de conversão) e, nos casos em que pudesse haver interferência com palavras homógrafas estrangeiras, as substituições seriam feitas de modo supervisionado. Para o efeito, descarregámos o *corpus* português e criámos uma lista com todas as palavras do *corpus* ordenadas por frequência.

Num primeiro momento, usámos os dois conversores na lista de palavras do *corpus* e num segundo confirmámos as conversões sugeridas pelos conversores com a ajuda do corretor ortográfico FLiP. Verificámos todos os termos individualmente até à frequência de 5 e demos a lista à equipa informática que gere as memórias de tradução Euramis para a substituição ser feita nas memórias de todas as instituições.

Alguns dias mais tarde demo-nos conta de que tinha havido problemas com a substituição quando vimos formas como «pacto» aparecerem agora como «pato». Uma rápida verificação permitiu ver que a equipa informática tinha substituído sequências de carateres e não palavras, contrariamente ao que havia sido indicado. Como da lista de palavras constavam substituições como *acto→ato* e *acta→ata*, a substituição de sequências acabou por fazer conversões como *pacto→pato* ou *impacto→impato,* o que, obviamente, não se pretendia.

O processo de substituição foi imediatamente interrompido e foi reposto o conteúdo da última cópia de segurança das memórias com a situação antes das substituições. Estudaram-se ainda alternativas técnicas para efetuar as substituições, mas o esforço estimado pela equipa que gere as memórias Euramis foi sempre julgado incomportável face aos benefícios e por aqui ficou a história da conversão das memórias das instituições europeias, que, até ao presente, nunca chegaram a ser convertidas.

A tradução automática estatística

Poderíamos ter ficado por aqui, mas tínhamos um novo problema para resolver: a **normalização** dos *corpora* da **tradução automática** (TA), que agora continham palavras AO45 e AO90 (ação/acção; diretiva/directiva; acto/ato, etc.). Tendo feito o trabalho de sapa, o DLP pediu ajuda informal a informáticos para desenvolver uma ferramenta — a base do **Orto90** — que pudesse usar a lista de conversão de forma otimizada. Isso foi mesmo feito e, desde 2012, a ferramenta tem desempenhado essas funções a contento.

A aplicação do AO90 aos *corpora* de treino em língua portuguesa do **serviço MT@EC** é incluída num **módulo de normalização** mais geral que também aplica uma boa parte das regras formais (aspas, travessões, separadores, carateres específicos para graus e ordinais, espaços protegidos, etc.) contidas no *Código de Redação Interinstitucional* — exemplos: *nº, n° → n.º*; *1a, 1ª → 1.ª*; *1o, 1º → 1.º*; *m2 →* m^2; *m3 →* m^3.

Embora tenhamos podido conviver há já cinco anos com memórias de tradução híbridas no que ao AO90 diz respeito, era impensável não normalizar os *corpora* de treino da TA, que ainda são na sua maioria pré-AO90. As palavras mais recorrentes afetadas pelo AO90 (diretiva, atividade, proteção) aparecem ainda nas memórias com 2/3 das ocorrências sob a forma AO45. Se não se normalizasse o *corpus*, teríamos resultados inferiores (em algumas experiências havia uma diferença de 5-6 pontos BLEU) e obteríamos uma mistura de AO45/AO90 com alguns brasileirismos, aspas e travessões de vários feitios, etc., fazendo recair no tradutor o esforço de atualizar a grafia nos casos em que a estatística da TA determinasse palavras com a ortografia AO45. É este atualmente o comportamento do *Google Translate* e do *Bing* da Microsoft, em que a ortografia AO45/AO90, português europeu/português do Brasil não é objeto de normalização e o resultado ortográfico é aleatório, por vezes até na mesma frase, como se pode ver nas figuras 1 e 2 a seguir[8].

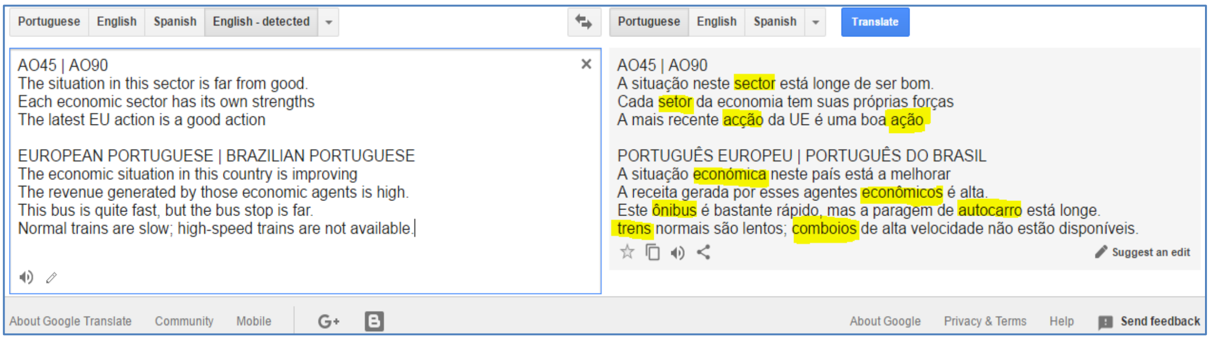


Figura 1: variação ortográfica AO45/AO90 e português europeu/português do Brasil no Google Translate

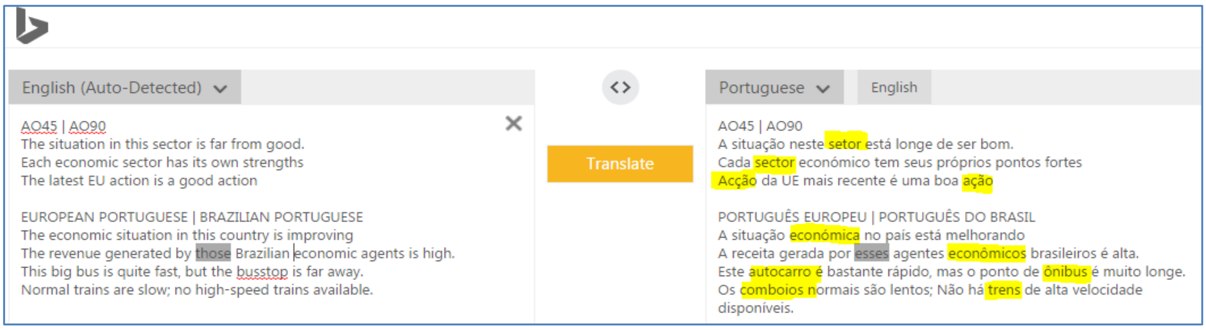


Figura 2: variação ortográfica AO45/AO90 e português europeu/português do Brasil no Microsoft Translator

Ora, se para efeitos de compreensão estes fenómenos não têm grande importância, se queremos produzir tradução automática o mais próxima possível da qualidade de publicação e não perder tempo a uniformizar, temos de aplicar as convenções ortográficas (e outras) de uma forma uniforme.

O trabalho com o AO90 aplicado à TA tem permitido obter um resultado ortográfico imediato atualizado e coerente e, ao mesmo tempo, caminhar para uma maior uniformização não só da aplicação das regras do AO90, mas de outras convenções, servindo de instrumento normalizador, a exemplo do que faz o módulo de normalização das regras formais de escrita adotadas pelas instituições. Ainda não perdemos a esperança de no futuro, quando as condições técnicas o permitirem, aplicarmos o **Orto90** às memórias de tradução, mas neste momento as nossas baterias estão assestadas a conseguir que a lista de conversão seja o mais exaustiva possível e em resolver e/ou mitigar os problemas de aplicação do AO90 dos corretores ortográficos que utilizamos.

Entretanto, entre 2012 e 2016 o tamanho do *corpus* de treino português duplicou, por ação da atividade tradutória, mas, sobretudo, pela incorporação de novos textos. Isto torna necessário atualizar a lista de substituições em função do novo *corpus* e verificar se tem havido problemas com a aplicação do AO90. É o que começámos a fazer, tendo já identificado vários milhares de novas entradas e alguns novos problemas.

A título de curiosidade, a situação em 2016 nas memórias de tradução para as cinco palavras mais comuns afetadas pelo AO90 é exemplificativa de como a nova ortografia ainda está longe de predominar nas memórias de tradução:

directiva (1 318 269), diretiva (448 809)	**25,4%** de 1 767 078 ocorrências
actividades (460 552), atividades (227 476)	**33,1%** de 688 028 ocorrências
protecção (458 941), proteção (229 454)	**33,3%** de 688 395 ocorrências
sector (376 625), setor (159 504)	**29,8%** de 536 129 ocorrências
acção (372 426), ação (165 587)	**30,8%** de 538 013 ocorrências

Como se pode verificar por esta pequena amostra, o AO90 ainda é muito minoritário nas memórias das instituições europeias.

Substituir o quê?

A estratégia de construção da **lista de conversão** foi delineada no artigo «As memórias de tradução e a ortografia»(9):

- **1.ª sublista** — **conversão**: palavras que mudam com o AO90 (exemplo: *directiva* passa a *diretiva*);
- **2.ª sublista** — **correção**: gralhas e erros independentes do AO90 (exemplo: **ímans* passa a *ímanes*);
- **3.ª sublista** — **normalização**: variantes ortográficas passíveis de harmonização (exemplo: *sobressalente* e *sobresselente*);
- **4.ª sublista** — **exclusão**: palavras que deveriam ser emendadas manualmente na memória Euramis.

<u>1.ª sublista — conversão</u>

São todas as palavras corretamente escritas com o AO45, alteradas pelos conversores AO45→AO90 — exemplos: *acção* → *ação*; *auto-estrada* → *autoestrada*; *carotenóides* → *carotenoides*.

É necessário ter em atenção dois casos particulares:

- Locuções, geralmente substantivas e adjetivas, em que o VOP[10] prevê a supressão dos hífenes, não sendo seguido por outros vocabulários, dicionários e conversores ou corretores ortográficos — exemplos: *caminho-de-ferro* → *caminho de ferro*; *mão-de-obra* → *mão de obra*.
- Novas duplas grafias do português de Portugal em que as predefinições dos conversores e corretores ortográficos implicam alteração da ortografia — exemplos: *carácter* → *caráter*; *espectro* → *espetro*. Notar que nalguns casos as predefinições variam entre ferramentas de conversão[11].

Teoricamente bastaria ficar por aqui…

2.ª sublista — correção

São as palavras com erros e gralhas (detetáveis pelos corretores) e que, em princípio, não são alteradas pelos conversores AO45→AO90, pois não respeitam as regras do Acordo Ortográfico de 1945, mas que têm que ser corrigidas. Alguns tipos de erros e gralhas:

- Erros resultantes do desconhecimento da complexidade das regras ortográficas da prefixação do AO45[12] — exemplos: **bio-resíduos* → *biorresíduos*; **bio-combustível* → *biocombustível*.
- Outros erros ortográficos e gralhas (**calendragem* → *calandragem*; **eligibilidade* → *elegibilidade*; **Pirinéus* → *Pirenéus*; **percursor* → *precursor*; **contéudo* → *conteúdo.*
- Decalques de regras ortográficas do inglês — exemplos: **ex-ante* → *ex ante*; *ex-post* → *ex post*
- Formas que estando corretas noutros contextos, não o estão nos nossos contextos — exemplos: *sedeada*[13] → *sediada*; *principio*[14] → *princípio*.
- Problemas ortográficos geralmente decorrentes de manipulações informáticas de texto, como:
  - Hifenações duplas ou não padrão — exemplos: *programa--quadro* → *programa-quadro*; *contra--ordenação* → *contraordenação*; *advogado-geral* → *advogado-geral*)
  - Palavras com carateres ricos corrompidos — exemplos: *indicacoes* → *indicações*; *dimensőes* → *dimensões*; *alteraçöes* → *alterações*)
  - Palavras coladas (*solicitaque* → *solicita que*; *EstadosMembros* → *Estados-Membros*)
- Grafias brasileiras (a DGT recebe textos originais em português do Brasil) — que não sendo erros propriamente ditos, têm de ser corrigidas antes da reutilização como português de Portugal — exemplos: *econômico* → *económico*; *indenizar* → *indemnizar*.

Posteriormente a 2012 surgiram novos casos, resultantes da má aplicação do AO90 ao português de Portugal:

- Novas duplas grafias do português de Portugal não previstas nas predefinições do corretor ortográfico — exemplo: *caraterística* → *característica.*
- «Hipercorreções» pós-AO90, com eliminação de consoantes pronunciadas no português de Portugal — exemplos: *seção* → *secção*; *fato* → *facto*; *contato* → *contacto*; *exeto* → *exceto.*

Poderíamos ficar por aqui, mas já que também estamos a trabalhar para a tradução automática estatística…

3.ª sublista — normalização

São variantes ortográficas já herdadas do AO45 e ortografias estrangeiras e passíveis de harmonização (a fixar igualmente na base IATE):

- Supressão de variantes independentes do AO90 — exemplos: *extractível*, *extractável* → *extraível*; *loiça* → *louça*; *lignite*, *lignito*, *linhito*, *linhita*, *linhite* → *lenhite*; *manganés*, *manganésio* → *manganês*; *radão* → *rádon*.
- Fixação de topónimos e outros nomes próprios estrangeiros em que se oscila entre diferentes grafias (muitas vezes francesa ou inglesa) — exemplos: *Al-Qaida*, *Al-Qaeda* → *Alcaida*; *Daesh*, *Daech* → *Daexe*; *Bahrain*, *Bahrein* → *Barém*; *Qatar* → *Catar*; *Machereque*, *Machrek*, *Mashrek* → *Maxerreque; Amsterdão* → *Amesterdão*.
- Ultrapassagem do conservadorismo ortográfico, que retarda a generalização das grafias portuguesas já consagradas nos dicionários de língua portuguesa(15) — exemplos: *hamburgers*, *hamburguers* → *hambúrgueres*; *pizzas* → *pizas*; *pellet* → *pélete*; *bulldozer* → *buldózer*; *tablet* → *táblete*; *jihadista* → *jiadista*.
- Fixação de grafias inglesas — exemplo: *off-shore* → *offshore*; *taskforce*, *task-force* → *task force*; *organization* → *organisation*.

4.ª sublista — exclusão

Palavras que deveriam ser **emendadas manualmente** na memória Euramis, pois são homógrafas de palavras estrangeiras frequentes nos nossos textos — exemplo: *sector*. No entanto, aplicando-se a lista às memórias da tradução automática, a sobrecorreção destas palavras não é estatisticamente um problema — exemplo: é preferível ter de corrigir pontualmente *setor* para *sector* numa expressão deixada em inglês, do que ter de corrigir repetidamente *sector* para *setor* numa frase redigida em português.

Uma conclusão prévia

Este trabalho de análise das listas de termos ensinou-nos e continua a ensinar-nos muito sobre a forma como certas **regras do AO45** eram desconhecidas e como outras **regras do AO90** são mal interpretadas; como o **original influencia a tradução**, pelo menos no aspeto ortográfico, e como há importantes problemas com a não utilização adequada do **corretor ortográfico** nos textos que acabam nas memórias Euramis.

Substituir como?

Optámos por só incluir na lista de conversão sequências sem espaços, tendo resolvido alguns casos conhecidos com espaços espúrios (p. ex., *advogado -geral*) diretamente na base de dados que contém as memórias (Euramis). A primeira razão é a própria metodologia da lista de frequência; a segunda é a baixa ocorrência desses fenómenos; a terceira, a simplicidade. Isto não obsta a que no futuro não possamos debruçar-nos sobre expressões multipalavras que queiramos normalizar.

Para evitar e/ou minimizar danos colaterais, a lista impõe sempre uma grafia aceitável. Por exemplo: ação, Ação, AÇÃO, mas Amesterdão, AMESTERDÃO e não amesterdão. Assim, a lista contém tantas variações de caixa quantas as aceitáveis no nosso *corpus*.

A ferramenta **Orto90** é composta por uma lista de palavras no formato **errado|certo** e por uma rotina informática programada em linguagem Perl. Quando se executa a rotina, pode incluir-se uma etapa de toquenização prévia do documento/*corpus* a processar e destoquenização depois de efetuadas as substituições. Os melhores resultados foram obtidos com duas passagens, a primeira sem toquenização — utilizando os espaços e uma série de sinais de pontuação como separadores — e a segunda passagem com toquenização/destoquenização utilizando o separador de palavras da linguagem de programação. No seu estado atual, trata-se de uma ferramenta para ser usada na linha de comando em Linux. Assim, dado um ficheiro de texto em UTF-8, a ferramenta faz uma pesquisa/substituição de cada uma das entradas da lista. Podemos disponibilizá-la a quem no-la solicitar no seu estado atual, mas a intenção é publicá-la com uma licença aberta logo que tenhamos a lista de entradas atualizada com o *corpus* 2016(16). É importante frisar que esta ferramenta converte, corrige e normaliza em função do *corpus* das memórias da União Europeia e das regras aplicadas pelas instituições e que se

alguém a quiser utilizar noutro contexto poderá ter de personalizar e/ou atualizar a lista de conversão em função do *corpus* que queira processar e das opções que queira aplicar.

Futuro

Uma vez terminado o trabalho de atualização da lista de conversão, o primeiro objetivo será consolidar as diversas componentes informáticas para poder proceder a atualizações regulares e publicar a ferramenta. É também nossa intenção fazer uma «fotografia» das palavras do *corpus* atual para comparar com a fotografia tirada no momento da próxima atualização, de molde a só analisar as palavras entretanto acrescentadas.

Uma das fontes de erros está associada à utilização do corretor ortográfico — atualmente o pacote **FLiP v.8**. Desde o início que se sabe que o corretor do FLiP não se alinha pelo VOP, nomeadamente em relação à aplicação das regras do hífen em locuções com preposições (exemplos: FLiP — caminhos-de-ferro, mão-de-obra; VOP — caminhos de ferro, mão de obra). É necessário um cuidado redobrado com estas palavras/locuções e, em caso de dúvida, consultar o VOP.

Como com qualquer corretor ortográfico generalista, o corretor FLiP desconhece muitas palavras frequentes nos nossos textos (ciberarmamento, fásico, macroprudencial, negacionismo, parajudicial, redifusão, etc.). Mas, mais misteriosamente — por vezes, mas nem sempre, nuns documentos sim, noutros documentos não, nuns computadores sim, noutros computadores não —, o corretor prescreve soluções contrárias ao acordo ou incoerentes com as opções definidas (atividade, mas extra-a**ct**ividade; setorial, mas plurisse**ct**orial, interse**ct**orial; prospetiva, mas prospe**ct**ivamente; afetar, mas reafe**ct**ar, reafe**ct**ação; fraturar, mas fra**ct**uração; etc.) ou deixa em paz duplas ortografias (ja**c**to, jato; Egi**p**to, Egito; telespe**c**tadores, telespetadores, etc.). Segundo a Priberam, a editora do FLiP, estas instabilidades com o corretor ortográfico resultariam, geralmente, de interferências com o programa de processamento de texto (Word). Por vezes, ao abrir-se um documento ou ao copiar-se um texto para um documento, o Word não assinala eventuais erros ou gralhas que nele possam existir, o que requer cuidado redobrado e a reativação regular do corretor.

Idealmente, as instituições e a administração pública portuguesa deveriam dispor de um **corretor ortográfico específico** que aplicasse as regras do AO90 de acordo com o VOP, as escolhas das duplas grafias, reconhecesse a esmagadora maioria do vocabulário jurídico e administrativo e permitisse uma rápida atualização centralizada das palavras e/ou regras.

Também seria interessante dispor de um **módulo gramatical** robusto que detete os problemas gramaticais e seja capaz, por exemplo, de identificar erros insidiosos como *foram*/*forma*, regências, construções bastardas como quer/ou* e seja/ou*. Algumas **funcionalidades de garantia da qualidade**, como uma lista negra de termos banidos/perigosos (calão, termos caídos em desuso, termos fortemente conotados) e/ou impróprios (República Libanesa e não República do Líbano; Federação da Rússia e não Federação Russa). Se quiséssemos ir mais longe nesta vertente, a ferramenta aplicaria as regras contidas no *Código de Redação Interinstitucional* e no *Guia do Tradutor* da DGT no que respeita a fórmulas, uso do espaço protegido, separador de milhares e das décimas e outros formalismos, na esteira do que faz a ferramenta de normalização usada no serviço MT@EC a que já se aludiu antes.

No entretanto…

No entretanto, podemos já tirar partido da lista de conversão do **Orto90** no nosso ambiente Windows e criar:

- um **dicionário suplementar comum** (CUSTOM.dic) com todas as palavras que o corretor não conhece; isto já foi feito(17), embora tenhamos tido alguns problemas com a distribuição centralizada;

- um **dicionário de exclusões** (ExcludeDictionaryPT0816.lex) para forçarmos o corretor ortográfico a:
  - só reconhecer a variante que queremos aplicar (*jato* e não *jacto*; *setor* e não *sector*); e
  - neutralizar soluções contrárias às fontes seguidas pelas instituições da UE (*caminho de ferro* e não *caminho-de-ferro*; *anti-dumping* e não *antidumping*).

No dicionário de exclusões poderão acrescentar-se:

- Palavras em que o corretor se afasta do VOP — exemplos: *caminho-de-ferro* → *caminho de ferro*; *mão-de-obra* → *mão de obra*, *não-proliferação* → *não proliferação*
- Palavras corretas mas com outro sentido de ocorrência muito pouco provável nos nossos textos — exemplos: *sedeada* → *sediada*; *principio* → *princípio*
- Variantes ortográficas reconhecidas pelo corretor mas passíveis de harmonização (incluindo palavras inglesas/americanas) — exemplos: *lignite*, *lignito*, *lignita*, *linhite*, *linhito*, *linhita* → *lenhite; antidumping* → *anti-dumping*

Conclusões

A primeira operação de conversão das memórias de tradução das instituições da União Europeia para o AO90 foi uma operação malograda. Já os trabalhos que levaram à criação do Orto90 foram extremamente profícuos porque permitiram:

a) adquirir um conhecimento muito mais aprofundado das questões de ortografia nos documentos da UE;
b) reparar uma série de problemas ortográficos cuja existência ou dimensão se desconhecia;
c) normalizar e estabilizar a ortografia de um vasto conjunto de termos; e
d) pôr à disposição da tradução automática um instrumento de normalização de todo o *corpus* em língua portuguesa das memórias Euramis (atualmente, aproximadamente 40 milhões de segmentos) utilizadas no treino dos motores da tradução automática da UE.

A tradução automática está, assim, transformada numa força de normalização (ortográfica e não só) dos textos em língua portuguesa produzidos pela DGT e pelos outros órgãos e instituições da UE que a utilizam no dia a dia. A base terminológica IATE procura alinhar-se pelo Orto90 amplificando, consolidando e, se necessário, justificando as decisões normalizadoras. Finalmente, o trabalho vai ter também eco na configuração dos corretores ortográficos de modo a evitar tanto quanto possível incoerências logo na fase de redação dos textos.

Hilario.Fontes@ec.europa.eu
Paulo.Correia@ec.europa.eu

---

(1) Wikipedia, *European Advanced Multilingual Information System*, https://de.wikipedia.org/wiki/European_Advanced_Multilingual_Information_System.
(2) Resolução do Conselho de Ministros n.º 8/2011 que determina a aplicação do Acordo Ortográfico da Língua Portuguesa no sistema educativo no ano letivo de 2011-2012, *Diário da República*, I série, n.º 17, de 25 de janeiro de 2011, https://dre.pt/application/dir/pdf1sdip/2011/01/01700/0048800489.pdf.
(3) Portal da Língua Portuguesa, http://www.portaldalinguaportuguesa.org.
(4) Comissão Europeia, *Código de Redação Interinstitucional*, http://publications.europa.eu/code/pt/pt-000100.htm.
(5) Comissão Europeia, *Código de Redação Interinstitucional*: 10.4. Pontuação, http://publications.europa.eu/code/pt/pt-4100412pt.htm.
(6) Comissão Europeia, Guia do Tradutor do Departamento de Língua Portuguesa, https://ec.europa.eu/info/files/portuguese-resources-portuguese-style-guide-dg-translation-house-style-guide_en.
(7) Correia, P., «Uso e abuso do hífen — relato de uma experiência»; Equipa linguística do Departamento de Língua Portuguesa, «Vocabulários do Acordo Ortográfico de 1990» *in* «a folha», n.º 32 — primavera de 2010, http://ec.europa.eu/translation/portuguese/magazine/documents/folha32_pt.pdf.
Correia, P., «Duplas grafias» *in* «a folha», n.º 33 — verão de 2010, http://ec.europa.eu/translation/portuguese/magazine/documents/folha33_pt.pdf.
Correia, P., Santos, V.N., «Duplas grafias no espaço da lusofonia» *in* «a folha», n.º 34 — outono de 2010, http://ec.europa.eu/translation/portuguese/magazine/documents/folha34_pt.pdf.

Equipa Linguística do Departamento de Língua Portuguesa, «Vocabulário das memórias de tradução» *in* «a folha», n.º 35 — primavera de 2011, http://ec.europa.eu/translation/portuguese/magazine/documents/folha35_vocabulario_pt.pdf.
Equipa Linguística do Departamento de Língua Portuguesa, «Conversores ortográficos e vocabulário das memórias de tradução» *in* «a folha», n.º 36 — verão de 2011,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha36_pt.pdf.
Fontes, H.L., Correia, P., «As memórias de tradução e a ortografia» *in* «a folha», n.º 37 — outono de 2011,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha37_pt.pdf.
Correia, P., «Nova ortografia — a experiência da DGT» *in* «a folha», n.º 39 — verão de 2011,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha39_pt.pdf.
(8) Exemplos recolhidos em 9 de outubro de 2016.
(9) Fontes, H.L., Correia, P., «As memórias de tradução e a ortografia» *in* «a folha», n.º 37 — outono de 2011,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha37_pt.pdf.
(10) Portal da Língua Portuguesa, *Vocabulário de Mudança*,
http://www.portaldalinguaportuguesa.org/novoacordo.php?action=novoacordo&act=list&version=pe.
(11) Correia, P., «Duplas grafias» *in* «a folha», n.º 33 — verão de 2010,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha33_pt.pdf.
(12) Correia, P., «Uso e abuso do hífen — relato de uma experiência» *in* «a folha», n.º 32 — primavera de 2010,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha32_pt.pdf.
(13) A forma existe (limpa com seda), mas nos nossos contextos o que se pretende dizer é «tem sede em».
(14) A forma existe (1.ª pessoa do presente do indicativo do verbo principiar), mas nos nossos contextos é muito raro utilizar-se a primeira pessoa do singular.
(15) Correia, P., «Palavras adaptadas ortograficamente» *in* «a folha», n.º 39 — verão de 2012,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha39_pt.pdf.
(16) Contactar os autores do artigo por correio eletrónico.
(17) Equipa linguística do Departamento de Língua Portuguesa, «Dicionário CUSTOM.DIC do Departamento de Língua Portuguesa — 2.ª versão (maio 2013)» *in* «a folha», n.º 41 — primavera de 2013,
http://ec.europa.eu/translation/portuguese/magazine/documents/folha41_dicionario_pt.pdf.

Notas sobre a ortografia dos quimbundismos e do quimbundo

Amarílis Pêgo
Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

O quicongo, a língua do antigo reino do Congo e do atual norte de Angola[1], foi a primeira língua banta de Angola com que os portugueses entraram em contacto. Porém, com a criação, mais a sul, da colónia de São Paulo da Assunção de Luanda, o **quimbundo**[2], a língua banta da cidade e da sua região interior, passou a ser a língua angolana com maior influência no português[3], fornecendo-lhe novas palavras para descrever novas realidades e uma nova cultura.

Estou-me **marimbando**[4] para o discurso do país de **tanga**[5].
Lá estão dois **cotas**[6] a fumar **cachimbo**[7].
Cambada[8]. Fiquei **banzado**[9] com a velocidade a que ele **bazou**[10].
Ficou **bué**[11] **ganzado**[12] depois de fumar a **liamba**[13] comprada na **candonga**[14].

Estava **xingando**[15] os **moleques**[16] do **quilombo**[17].
O **caçula**[18] tirou um **cochilo**[19] depois de estar fumando **maconha**[20].
Quero **cuiar**[21] você com muitos **cafunés**[22].
Moqueca[23] de camarão e abobrinha com talharim de **fubá**[24] de milho.

Estamos finalmente no **cacimbo**[25].
Comi **bué de jinguba**[26] com **jindungo**[27].
Os meus **cambas**[28] gostam de dançar **quizomba**[29] e **semba**[30].
Mexe essas **quinamas**[31]!
Que grande **maca**[32]. O meu **cota** chegou mais cedo à casa. Portanto, mais ninguém vai **desbundar**[33].

A **cubanga**[34] desses dois não acaba. Chiça! Isto já me cheira a **fundanga**[35]!
Essa dama é **bué cambuta**[36]!
Essa **cuzanga**[37] toda tem de acabar. Fecha a torneira, se faz favor!
Ele **fez-lhe a boa muxima**[38] e foi autorizado a sair mais cedo.
Atenção aos **mujimbos**[39] da vizinha **zongola**[40]. Ela gosta muito de **zuelar**[41].

Quimbundismos e ortografia

As palavras assinaladas a negrito nas frases anteriores são **quimbundismos**[42] — termos introduzidos no léxico português a partir do quimbundo. Da música à política, passando pelo vocabulário mais ou menos informal de jovens e menos jovens, com um cheirinho a África, a Brasil ou mesmo sem cheiro nenhum, antigos ou recentes, largas dezenas de quimbundismos estão entre nós e são uma das particularidades das variedades da língua portuguesa faladas no triângulo Angola-Brasil-Portugal.

> Naturalmente, no sentido oposto, também o quimbundo foi enriquecido no contacto com o português, incorporando as novas palavras na lógica da língua quimbunda. Exemplos: *divulu* (de livro) / *madivulu* (livros); *ngalafa* (de garrafa) / *kangalafa* (garrafinha); *kujandala* (jantar).

Sendo palavras portuguesas, os quimbundismos seguem a 100% as **regras ortográficas e morfológicas do português** e não as da língua de origem, o quimbundo. Convém deixar bem claro que uma coisa são palavras portuguesas com origem no quimbundo, outra são palavras quimbundas, escritas com regras ortográficas do quimbundo e para as quais — tal como se faz para as palavras escritas em inglês, espanhol ou francês — deverá utilizar-se o itálico nos textos em português[43].

> Exemplos: banto, mas *bantu*; quimbundo, mas *kimbundu*; missangas, mas *misanga*.

Dentro das regras ortográficas próprias do português, detetam-se alguns padrões na ortografia dos quimbundismos registados em vocabulários e dicionários portugueses e brasileiros e algumas duplas grafias. Por exemplo, nos quimbundismos, **j**, **ss** e **x** são (quase) sempre utilizados em vez de *g*, *ç* e *ch* para as transcrições dos sons /ʒ/, /s/ ou /ʃ/[44].

- Para o som /ʒ/, **j(e/i)** (em vez de *g(e/i)*) — **jindungo**, e não *gindungo, **canjica**, e não *cangica. Por exemplo, a atual grafia portuguesa de **Malanje** é uma evolução de Malange (não é de excluir aqui a possível influência das regras ortográficas do quimbundo).
- Para o som /s/ intervocálico, **ss** (em vez de *ç*) — a prática lexicográfica luso-angolana (**ss**)[45] diverge da prática brasileira (**ç**). Veja-se: **missanga** (pt-PT/AO) e **miçanga** (pt-BR); **caçula**, quimbundismo chegado via Brasil; **musseque**, chegado diretamente de Angola; etc. Embora evitável, esta divergência tem a vantagem de indicar de que braço do triângulo Angola-Brasil-Portugal o quimbundismo chegou a Portugal.
- Para o som /ʃ/, **x** (em vez de *ch*, ou mesmo do *sh* inglês) — xingar, e não *chingar, Moxico e não *Mochico.

<u>k, w, y</u>

Pela mesma lógica, deve evitar-se usar mais do que uma letra para um mesmo som. Utilizam-se nas palavras portuguesas com origem no quimbundo **c(a,o,u)** e **qu(e,i)** para o som /k/, **u** para o som /w/ e **i** para o som /y/. Nos países africanos de língua oficial portuguesa existem, de facto, muitas palavras que se escrevem com *k*, *w* e *y*, mas são palavras das línguas nacionais, escritas com os respetivos alfabetos.

> Apesar da inclusão no alfabeto das letras *k*, *w* e *y*, mantiveram-se, no entanto, as regras já fixadas anteriormente [Acordo Ortográfico de 1945[46]] quanto ao seu uso restritivo, pois existem outros grafemas com o mesmo valor fónico daqueles. Se, de facto, se abolisse o uso restritivo daquelas letras, introduzir-se-ia no sistema ortográfico do português mais um fator de perturbação, ou seja, a possibilidade de se representar, indiscriminadamente, por aquelas letras fonemas que já são transcritos por outras.[47]

Devem evitar-se, assim, estas três letras nas palavras portuguesas de origem quimbunda. Ao contrário de outros dicionários, nem sempre é essa a prática dos dicionários de língua portuguesa da Porto Editora (Infopédia) e Plural Editores(48), fortemente implantados no mercado angolano e que têm registado muito do vocabulário típico do **português angolano**, que utilizam ortografias quimbundas sem itálico, sobretudo no que se refere ao uso do *k*. Veja-se o exemplo da forma como a palavra quizomba é registada nos dicionários eletrónicos de acesso livre:

- Infopédia: kizomba(49); quizomba (ver: kizomba)
- Priberam: quizomba(50); *kizomba* (palavra quimbunda)
- Aulete: quizomba(51)
- Michaelis: quizomba(52)

Nos quimbundismos, por analogia com as palavras de origem latina, utiliza-se o **o** final para representar o som /u/ final átono.

> Não é lícito o emprego do *u* final átono em palavras de origem latina. Escreve-se, por isso: moto, em vez de mótu (por exemplo, na expressão de moto próprio); tribo, em vez de tríbu(53). Exemplos: quimbundo e não quimbúndu.

(n)g

A pré-nasalação de consoantes, própria das línguas bantas, transforma-se nos quimbundismos na nasalação das vogais anteriores ou, no início das palavras, na junção de vogais nasaladas (prótese) ou na supressão pura e simples da pré-nasalação (aférese). Exemplos:

adição de sons:	Ngola	prótese — Angola
supressão de sons:	*nganza*	aférese — ganza

Os quimbundismos, ao entrarem na língua portuguesa, comportam-se também morfologicamente como palavras portuguesas, passando, por exemplo, os verbos a conjugar-se como verbos portugueses do primeiro grupo ou as palavras a formar o plural com *s* final(54). Ou, citando Cláudio Moreno, sobre a questão dos plurais irregulares em palavras oriundas de outras línguas, decalcando as regras dessas mesmas línguas:

> Julgo, entretanto, que imitar essa prática no Português seria criar uma injustificável exceção ao paradigma (imaginem «os talibã»!) e ignorar a extraordinária capacidade que nosso idioma tem de deglutir os vocábulos estrangeiros e nacionalizá-los fonológica, ortográfica e morfologicamente. Já escrevi várias vezes sobre isso: para entrar no Português, o vocábulo estrangeiro tem de aprender a dançar miudinho, tratando de comportar-se como seus colegas nativos. Um talismã, dois talismãs; um talibã, dois talibãs.(55)

pt	g.	pref.	kmb	etimologia	sinónimo
jinguba(s)	f.	***ji-***	*jinguba*	**plural** de *nguba*, amendoim	amendoins
mabaça(s)	m./f.	***ma-***	*mabasa*	**plural** de *kabasa*, gémeo, irmão gémeo	gémeo/a(s)
missanga(s) (PT) miçanga(s) (BR)	f.	***mi-***	*misanga*	**plural** de *musanga*, conta de vidro	
musseque(s)	m.	***mu-***	*museke*	**lugar** de *seke*, areia	favela(s)
candonga	f.	***ka-***	*kandenge*	**dimin.** de *ndenge*, menor, negócio pequeno	mercado negro
quimbundo	m.	***ki-***(56)	*kimbundu*	**língua** dos *mbundu*, pretos	
curinga(s)	m.	***ku-***	*kuringa*	**infinitivo** do verbo *kuringa*, fingir	jóquer(es)

Algumas questões ortográficas e morfológicas do quimbundo

A Companhia de Jesus, nas suas tarefas de evangelização, cedo foi levada a estudar as línguas dos povos com que os portugueses entraram progressivamente em contacto, e as atuais línguas nacionais de Angola não foram exceção. O quicongo(57) e o quimbundo foram, por esta ordem, as primeiras

línguas bantas a serem escritas com o alfabeto latino, no já distante século XVII — notando-se, porém, que o suaíli já era escrito, com o alfabeto árabe, desde o século XI.

Em 1697, foi publicada em Lisboa a obra *A Arte da Língua de Angola*(58), em que o padre Pedro Dias, baseando-se nas anotações do catecismo de 1641 do padre Francisco Pacconio(59), lança as bases de uma ortografia/transcrição do quimbundo modelada no latim e no português, que foi sendo afinada por outros autores, principalmente nos séculos XIX e XX. Destacam-se algumas obras: em 1889, *Kimbundu grammar — gramática elementar do kimbundu ou língua de Angola*(60), de Héli Chatelain, missionário protestante suíço; em 1893, o *Ensaio de dicionário kimbúndu-português*, de Joaquim Dias Cordeiro da Mata(61); em 1934, a *Gramática de Kimbundo* de José Luís Quintão(62). Nestas obras é introduzido o *k* na ortografia do quimbundo, aproximando-se mais da ortografia banta das regiões sob controlo inglês, embora mantendo outras soluções próximas da ortografia portuguesa. Veja-se o exemplo da oração do Pai Nosso:

1641: **Catecismo de Pacconio**(63)	1889: **Kimbundu Grammar**(64)	**Vaticano**(65)
Tatetu, uecala co maulu, acondeque o rigina riae, quize co tuecala o quifuchi quiae, acuzelese o muchima uae inga boxi inga beulu.	*Tat'etu uala ku maulu, axile o rijina rié, kize ku tuala o kifuxi kié, a ku zelese o muxima ué boxi kala bulu.*	Pai Nosso que estais nos céus, santificado seja o vosso Nome, venha a nós o vosso Reino, seja feita a vossa vontade assim na terra como no céu.
O mussa uetu ua izua yosso tube o rierino; tuequie o macongo etu nguequi tuequia anha aturia macongo. Cu tu equie pe curigia mu quituxi, tubanguele bó mu quiaiba.	*O musa uetu ua izúa ioso tu be-u lelu; tu ehele o makongo kala ki tu (m')eha aná a tu ria makongo. Ku tu ehe pe ku ri bala mu kituxi, (maji) tu bangéle mu kiaiiba.*	O pão nosso de cada dia nos dai hoje, perdoai-nos as nossas ofensas assim como nós perdoamos a quem nos tem ofendido, e não nos deixeis cair em tentação, mas livrai-nos do Mal.

Na segunda metade do século XX, na sequência da descolonização, a Organização da Unidade Africana encorajou a promoção das línguas africanas e a alfabetização das populações. São também de destacar os estudos enquadrados pela UNESCO e Associação Fonética Internacional (AFI) para a criação de um alfabeto africano de referência e de regras para a fixação das transcrições de uma língua e das línguas da mesma família. Visava-se a estandardização, unificação e harmonização de grafias dentro de um país e eventuais acordos ortográficos com países vizinhos que partilham a mesma língua(66).

No contexto deste movimento, o Instituto de Línguas Nacionais desenvolveu estudos para a reforma ortográfica do quimbundo (e de outras línguas bantas), na sequência dos quais as autoridades angolanas aprovaram a título experimental, pela Resolução n.º 3/87, de 13 de maio, do Conselho de Ministros(67), um novo **alfabeto do quimbundo**, que pode ser consultado em anexo a este artigo.

Foram também fixadas algumas regras ortográficas segundo as regras da AFI(68):

<u>1.ª regra:</u> Vogais: a, e, i, o, u
Semivogais: w, y
h aspirado
As palavras acabam sempre em vogal
<u>2.ª regra:</u> Não podem ocorrer duas vogais seguidas.
ua, ue, ui, uo, uu > wa, we, wi, wo, wu (ex.: Lwanda em vez de Luanda)
ia, ie, ii, io, iu > ya, ye, yi, yo, yu (ex.: Soyo em vez de Soio)
<u>3.ª regra:</u> Codificação com base em pronúncia não influenciada por outras línguas
<u>4.ª regra:</u> w antes de vogais e nunca no fim de palavras
<u>5.ª regra:</u> y antes de vogais e nunca no fim de palavras

Na introdução do seu *Dicionário de Etnologia Angolana*(69), Adriano Parreira chama a atenção para diferenças entre a ortografia do quimbundo (e de outras línguas bantas angolanas) e a ortografia do português.

> Na transcrição de palavras africanas, procurámos utilizar os símbolos do Alfabeto Fonético Internacional, razão pela qual os fonemas que em português correspondem a *qu*, *gu*, *ss* ou *ç*, são, nos vocábulos das línguas nacionais angolanas, respetivamente representados por *k*, *g* e *s*. As especificidades dos dialetos angolanos quanto à pronunciação da fricativa obrigam a ajustamentos necessários, como é o caso da consoante *x*, que corresponde ao som [ʃ] (…). (…) nas línguas nacionais angolanas ocorrem variações fonéticas similares, sejam elas a nível local, regional ou nacional, pelo que o valor fonético representado por x poderá, em alguns casos, variar entre [ʃ] e [tʃ]. No que se refere à nasalação [ɲ] ela é transcrita pelo conjunto das consoantes *ny*. (…)

Na prática, as igrejas tiveram e continuam a ter um papel importante na consolidação do quimbundo e da sua ortografia. As diferentes igrejas angolanas parecem ainda não ter aderido à reforma ortográfica do quimbundo proposta pelo Instituto de Línguas Nacionais, mantendo-se fiéis à ortografia de Chatelain e Quintão.

Matendelelu, 6:5-7(70) — com versão áudio	**Deuteronómio 6:5-7**(71)
5 *Zola-phe Jihova, Nzambi iê ni muxima uê uoso, ni muenhu uê uoso, ni nguzu iê ioso.* 6 *Maka u ngala ku ku tendelesa lelu'idi, u ma lunda ku muxima uê.* 7 *An'ê u a longa muene o ku a bhaka ni ku a kumbidila; o ki u xikama m'onzo iê, banza maka enhá; o ki u kala mu kuenda njila, enda ni maka enhá; o ki u zeka, zeka ni maka enhá; o ki u balumuka ku kilu, kuata maka enhá.*	5 Amarás, pois, ao Senhor teu Deus de todo o teu coração, de toda a tua alma e de todas as tuas forças. 6 E estas palavras, que hoje te ordeno, estarão no teu coração; 7 e as ensinarás a teus filhos, e delas falarás sentado em tua casa e andando pelo caminho, ao deitar-te e ao levantar-te.

Tendo em conta o que pode também observar-se nos textos angolanos disponíveis na Internet, abundam grafias várias do quimbundo, construídas de forma aparentemente natural pelos utilizadores da língua, que nelas refletem livremente as várias influências a que se encontram sujeitos, como seja a alfabetização em português (língua atualmente falada por cerca de 90% da população angolana), a ortografia tradicional de Chatelain e Quintão (com *k*, mas sem *w* ou *y*; com *nh*, mas sem *ss*) ou a interferência de outras ortografias bantas de países limítrofes.

> No nome da telenovela angolana ***Jikulumessu — Abre o Olho***(72), repare-se no *ss* dobrado, desconhecido na ortografia banta. Em rigor, em quimbundo a ortografia reformada seria Jikulume**su**(73), com um só *s*, o que poderia acarretar problemas na leitura da última sílaba.

Algumas particularidades do quimbundo (e de outras línguas bantas)

1. As letras *m* e *n* não nasalam a vogal antecedente, mas a consoante imediata. Exemplos:
   - *ambula* = a·mbu·la (e não am·bu·la)
   - *imvo* = i·mvo (e não im·vo)
   - *ndongo* = ndo·ngo (e não n·don·go)

2. Não há masculino e feminino.
3. Tal como nas restantes línguas bantas, o quimbundo utiliza prefixos nominais para indicar categorias de palavras, singulares e plurais. Ver quadro seguinte, adaptado da *Gramática de Kimbundo* de Quintão.

	kmb				**pt**	
	sing.		**pl.**		**sing./pl.**	
1	***mu-***	*muhatu*	***a-***	*ahatu*	mulher(es)	seres animados
2	***mu-***	*muxi*	***mi-***	*mixi,*	pau(s), árvore(s)	seres inanimados
3	***ki-***	*kinama*	***i-***	*inama*	perna(s)	instrumentos, etc.
4	***di-***	*dibengu*	***ma-***	*mabengu*	ratazana(s)	objetos de grandeza
5	***u-***	*uta*	***mau-***	*mauta*	arma(s)	termos abstratos
6	***lu-***	*lubambu*	***malu-***	*malubambu*	corrente(s)	objetos de extensão
7	***tu-***	*tubia*	***matu-***	*matubia*	fogo(s)	termos abstratos
8	***ku-***	*kufua*	***maku-***	*makufua*	morte(s)	termos verbais
9	***i-***	*imbua*	***ji-***	*jimbua*	cão/cães	animais
10	***ka-***	*kaditadi*	***tu-***	*tuditadi*	pedrinha(s)	diminutivos

Uma reflexão à margem

Sabendo-se que o /s/ intervocálico é representado nos alfabetos bantos por **s** e sabendo-se que o **m** antes de b e p e o **n** antes de d, j, t, v e z têm um papel de pré-nasalação da consoante seguinte, será talvez a ocasião de voltar a refletir sobre a ortografia de alguns exónimos da África subsariana inscritos no *Código de Redação Interinstitucional*: Lesoto ou Lessoto?; Maseru ou Massero[(74)]?; (/(əm)bɑˈbɑn(i)/) Mebabane ou Babane ou Ambabane ou Imbabane?

Amar.pego@gmail.com
Paulo.Correia@ec.europa.eu

Anexo: Alfabeto quimbundo

A, B, Bh, D, E, F, **Ng**, H, I, J, K, L, M, N, Ny, O, Ph, S, T, Th, U, V, W, X, Y, Z[(75)]

letra	fonética	equivalente português	palavra em quimbundo	«transliteração»
a	/a/ ~ /ɑ/	a (em *há*) a (em gl: irmán)	*kamba* (camarada)	camba
b	/b/	b (em *banto*)	*sambwadi* (sete)	sambuádi
bh	/bʰ/ ~ /v/		*lubhambu* (corrente)	lubambo
d	/d/	d (em *dendê*)	*dikwini* (dez)	diquíni
e	/e/ ~ /ɛ/	e (em *vê*) e (em *sé*)	*museke* (musseque)	musseque
f	/f/	f (em *faca*)	*fuba* (farinha)	fubá
ng	/ˆg/		Ngola (Angola)	Gola
h	/h/	h (em Holywood) rr lisboeta	*ihyi* (quê?)	irrii
i	/i/	i (em *vida*)	*wengi* (outro, diferente)	uêngui
j	/ʒ/	j (em *já*)	*jinguba* (amendoins)	jinguba
k	/k/	c (em *carro*)	*kubata* (casa modesta, simples)	cubata
l	/l/	l (em *lua*)	*lukwaku* (mão)	lucuaco
m	/m/	m (em *mão*)	*moxi* (um)	móxi
n	/n/	n (em *não*)	*samanu* (seis)	samano
ny	/ɲ/	nh (em *minhoca*)	*nyoka* (cobra)	nhoca
o	/o/ ~ /ɔ/	o (em *hoje*) o (em *só*)	*ososo* (qualquer)	ossosso
ph	/pʰ/ ~ /p/		*phi* (silêncio!)	pi
s	/s/ ~ /ts/	ss (em *massa*)	*misanga* (missangas, contas de vidro)	missanga
t	/t/	t (em *tomate*)	*ntanga* (pano)	ntanga
th	/tʰ/		*muthu* (pessoa)	mutu
u	/u/	u (em *caju*) o (em *banto*)	*tatu* (três)	tátu
v	/v/	v (em *vaca*)	*divwa* (nove)	dívua
w	/w/	u (em *quando*)	*kwanza* (moeda angolana)	cuanza
x	/ʃ/	x (em *xerife*)	*xala* (interjeição de despedida)	xala
y	/j/[(76)]	i (em *paio*)	*yadi* (dois)	iádi
z	/z/ ~ /dz/	z (em *zebra*)	*kizomba* (ritmo angolano)	quizomba

Fonte (letras e fonética): Quadro Geral dos Alfabetos em Línguas Nacionais, Angola 1980[(77)]

---

(1) Atuais províncias de Cabinda, Zaire e Uíje.
(2) Quimbundo, ou bundo do norte, a língua dos ambundos, atualmente a terceira língua banta mais falada em Angola, com cerca de quatro milhões de falantes no centro-norte de Angola, atuais províncias de Luanda, Bengo, Cuanza Norte, Malanje e parte de Cuanza Sul.
(3) Cf. Google (18.3.2017)
"do quimbundo" site:infopedia.pt — 1280
"do quicongo" site:infopedia.pt — 315
"do umbundo" site:infopedia.pt — 109
"do quioco" site:infopedia.pt — 20

"do cuanhama" site:infopedia.pt — 10

(4) Do quimbundo *ma-* (prefixo do plural) + *rimba*, «tambor». O *r* apareceu no quimbundo por contacto com o português.
(5) Do quimbundo *ntanga*, «pano».
(6) Do quimbundo *kota*, «superior».
(7) Do quimbundo *kixima*, «poço» ou do prefixo diminutivo *ka* + *humbu* «nome de um instrumento para fumar».
(8) Do quimbundo *kamba*, «camarada», + *-ada*.
(9) Do quimbundo *kubanza*, «pensar», «refletir».
(10) Do quimbundo *kubaza*, «romper».
(11) Do quimbundo *mbuwe*, «abundância», «fartura».
(12) Do quimbundo *nganza*, «cabaça».
(13) Do quimbundo *lyamba*, «cânhamo».
(14) Do quimbundo *ka-* (prefixo diminutivo) +*ndenge*, «pequeno; menor», isto é, «negócio pequeno; negócio insensato».
(15) Do quimbundo *xinga*, «injuriar».
(16) Do quimbundo *muleke*, «menino».
(17) Do quimbundo *kilombo*, «povoação».
(18) Do quimbundo *kazuli*, «o último da família».
(19) Do quimbundo *kuxila*, «dormitar», por derivação regressiva.
(20) Do quimbundo *makanya*, «tabaco», «erva-santa».
(21) Do quimbundo *kuya*, «bom».
(22) Do quimbundo *kafu'nu*, «cravar», «enterrar», afago suave na cabeça.
(23) Do quimbundo *mukeka*, guisado de peixe ou marisco, preparado com óleo de palma e leite de coco.
(24) Do quimbundo *fuba*, «farinha», farinha de milho ou de arroz usada em culinária.
(25) Do quimbundo *kixima*, «poço de água». Estação mais fria, com nevoeiro denso e húmido.
(26) Do quimbundo *jinguba*, plural de *nguba*, «amendoim».
(27) Do quimbundo *jindungu*, plural de *ndungu*, «amendoim», a partir de *kulungamena*, «agrupar-se», alusão à disposição.
(28) Do quimbundo *kamba*, «camarada».
(29) Do quimbundo *kizomba*, ritmo africano, a partir de *kuzomba*, «seguir pausadamente».
(30) Do quimbundo *semba*, «umbigada». Cf. Samba.
(31) Do quimbundo *kinama*, «perna».
(32) Do quimbundo *maka*, «palavra», problema.
(33) Do quimbundo *mbunda*, «nádegas», farrar.
(34) Do quimbundo *kubanga*, «disputa», «peleja».
(35) Do quimbundo *fundanga*, «pólvora», mau cheiro, fedor.
(36) Do quimbundo *kambuta*, «baixinho», a partir de *kubuta*, «ser baixo».
(37) Do quimbundo *kuzanga*, «esbanjar», esbanjamento.
(38) Cativar, lisonjear, ganhar confiança.
(39) Do quimbundo *mujimbo*, «boato», «zunzum».
(40) Do quimbundo *kuzongola*, «espiar», fofoqueira, faladora.
(41) Do quimbundo *kuzwela*, «dizer mal», «criticar».
(42) Os quimbundismos são um caso particular de bantoísmos — termos introduzidos no léxico português a partir de qualquer das línguas do grupo banto. Foram consultadas as notas etimológicas indicando a origem no quimbundo incluídas em vários dicionários portugueses e brasileiros, com destaque para a Infopédia.
(43) Outra atitude seria uma forma de discriminação: palavras inglesas seriam em itálico, enquanto palavras em quimbundo não teriam itálico, como se o quimbundo fosse uma variante do português e não uma língua de pleno direito.
(44) Estes padrões poder-se-ão comparar com os usados para a transcrição dos mesmos sons em vocábulos portugueses originários de línguas ágrafas (como o tupi-guarani) ou de línguas com alfabetos não latinos (como o árabe):

- Nos quimbundismos, contrariamente aos arabismos, privilegia-se o *j* em vez de *g*. Malanje, mas Algés, Almoçageme, Alfragide, etc.; jimbo, mas alforge (alforje pt-BR), beringela (berinjela pt-BR), etc.
- Contrariamente aos arabismos, nos quimbundismos luso-angolanos privilegia-se o *ss* em vez de *ç*. Cassinga, Quissama (cf. povoação de Quiçamã no Brasil), mas Almoçageme, Bucelas, Mucifal, Meleças, Odeceixe, Moçul, Faiçal, Moçambique, etc.; quissanje, mas açougue, açúcar, sargaço, etc.
- Neste caso (*x*), quimbundismos e arabismos seguem a mesma regra.

(45) Esta tradição parece ser extensível a outros topónimos da África subsariana — Bissau, Burquina Fasso, Iamussucro, Quinxassa, etc. —, com a exceção dos nomes de origem árabe do litoral do oceano Índico — Moçambique, Mombaça, etc.
(46) O Acordo Ortográfico de 1945, ainda em vigor em Angola, refere que: «O *k*, o *w* e o *y* mantêm-se nos vocábulos derivados eruditamente de nomes próprios estrangeiros que se escrevam com essas letras: *frankliniano, kantismo; darwinismo, wagneriano; byroniano, taylorista*. Não é lícito, portanto, em tais derivados, que o *k*, o *w* e o *y* sejam substituídos por letras vernáculas equivalentes: *cantismo, daruinismo, baironiano*, etc.», Portal da Língua Portuguesa, *Acordo Ortográfico de 1945*, http://www.portaldalinguaportuguesa.org/acordo.php?action=acordo&version=1945.
(47) Portal da Língua Portuguesa, *Acordo Ortográfico de 1990 — Anexo II da Legislação Portuguesa*, http://www.portaldalinguaportuguesa.org/acordo.php?action=acordo&version=1990b.
(48) Plural Editores África, *Dicionário Escolar da Língua Portuguesa*, 2015, http://www.pluraleditores.co.ao/catalogo/livraria/ficha/dicionario-escolar-da-lingua-portuguesa?id=171285.
(49) Infopédia, «kizomba», https://www.infopedia.pt/dicionarios/lingua-portuguesa/kizomba.
(50) Priberam, *Dicionário Priberam da Língua Portuguesa*, «quizomba», https://www.priberam.pt/dlpo/quizomba.

(51) Lexikon Editora Digital, *Dicionário Aulete Digital*, «quizomba», http://www.aulete.com.br/quizomba.

(52) Melhoramentos, *Michaelis Dicionário Brasileiro de Língua Portuguesa*, «quizomba», http://michaelis.uol.com.br/busca?r=0&f=0&t=0&palavra=quizomba.

(53) Portal da Língua Portuguesa, *Acordo Ortográfico da Língua Portuguesa de 1990*, http://www.portaldalinguaportuguesa.org/acordo.php?action=acordo&version=1990.

(54) Em sentido oposto, o facto de no Brasil haver a tendência, entre certos grupos e em certos registos, de não pronunciar o *s* final dos plurais será resultado da influência dos povos vindos de Angola, pois nas línguas bantas o plural é feito no início das palavras.

(55) Moreno, C., *Sua Língua*, «Talibã, talebã, taliban ou taleban», http://sualingua.com.br/2009/05/15/taliba-taleba-taliban-ou-taleban/.

(56) Noutras línguas bantas o prefixo de língua é xi, si, u, etc. Cf. «Línguas oficiais de países terceiros», *in* «a folha», n.º 35, http://ec.europa.eu/translation/portuguese/magazine/documents/folha35_pt.pdf.

(57) O padre jesuíta Mateus Cardoso editou em Lisboa, em 1624, o primeiro catecismo em português e congo.

(58) Internet Archive, Dias, P., *Arte da Língua de Angola*, Oficina Miguel Deslandes, Lisboa, 1697, https://archive.org/details/artedalinguadean00dias.

(59) Pacconio, F., *Gentio de Angola Suficientemente Instruído nos Mistérios de Nossa Santa Fé*, Domingos Lopes Rosa, Lisboa, 1642.

(60) Internet Archive, Chatelain, H., *Kimbundu Grammar: Grammatica Elementar do Kimbundu ou Lingua de Angola*, Charles Scuchardt, Genebra, 1888-89, https://archive.org/details/kimbundugrammar00chatgoog.

(61) Internet Archive, Mata, J. D. C. da (coord.), *Ensaio de Diccionario Kimbúndu-Portuguez*, António Maria Pereira, Lisboa, 1893, https://archive.org/details/ensaiodediccion00mattgoog.

(62) Quintão, J. L., *Gramática de Kimbundo*, Edições Descobrimentos, Luanda, 1934.

(63) Pacconio, F., *Gentio de Angola Suficientemente Instruído nos Mistérios de Nossa Santa Fé*, Domingos Lopes Rosa, Lisboa, 1642.

(64) Internet Archive, Chatelain, H., *Kimbundu Grammar: Grammatica Elementar do Kimbundu ou Lingua de Angola*, Charles Scuchardt, Genebra, 1888-89, https://archive.org/details/kimbundugrammar00chatgoog.

(65) Vaticano, *Catecismo da Igreja Católica*, «Quarta Parte: A Oração Cristã», http://www.vatican.va/archive/cathechism_po/index_new/p4s2_2759-2865_po.html.

(66) Não é o caso do quimbundo, língua limitada ao território angolano, mas é o caso de outras línguas nacionais de Angola, que também são faladas no Congo, República Democrática do Congo (antigo Zaire), Zâmbia e Namíbia.

(67) Embaixada de Angola em Portugal, *Línguas Nacionais*, http://www.embaixadadeangola.org/cultura/linguas/set_lnac.html.

(68) Cf. Augusto, M. A., *Morfologia Contrastiva entre Português e Kimbundu: Obstáculos e suas Causas na Escrita e Ensino da Língua Portuguesa entre os Kimbundu em Angola*, https://sapientia.pucsp.br/bitstream/handle/19178/2/Mois%C3%A9s%20Alves%20Augusto.pdf.

(69) Parreira, A., Dicionário de Etnologia Angolana, Plural Editores/Porto Editora, 2013, http://www.pluraleditores.co.ao/catalogo/livraria/ficha/dicionario-de-etnologia-angolana?id=14845215.

(70) Jimbangi ja Jihova, *Longa o Tuana Tuê*, «Dimatekenu», https://www.jw.org/kmb/Madivulu/madivulu/Longa-o-Tuana-Tu%C3%AA/dimatekenu/.

(71) União Sefardita Hispano-Portuguesa, *Bíblia — Versão Hebraica do Mekhon_Mamre*, «Deuteronômio», http://www.judaismo-iberico.org/interlinear/tanakh/0506PT.HTM.

(72) Telenovela angolana transmitida pela TPA1.
Semba, *Jikulumessu*, http://www.semba-c.com/pages/view/64/jikulumessu.

(73) Palavra formada por aglutinação de *jikula o mesu* ou *jikul'o mesu*.

(74) Lugar do arenito vermelho. Cf. musseque (lugar de areia).

(75) Não inclui C, G, P, Q, R.

(76) /y/ no original. No AFI, /y/ corresponde ao som *u* do francês.

(77) Wikimeida Commons, *File: Quadro General Dos Alfabetos Em Linguas Nacionais,Angola 1980.pdf*, https://commons.wikimedia.org/wiki/File:Quadro_General_Dos_Alfabetos_Em_Linguas_Nacionais,Angola_1980.pdf.

Letras do alfabeto hebraico e meses do calendário judaico

Paulo Correia
Jorge Madeira Mendes
Direção-Geral da Tradução — Comissão Europeia

בְּרֵאשִׁית, בָּרָא אֱלֹהִים, אֵת הַשָּׁמַיִם, וְאֵת הָאָרֶץ.
No princípio criou Deus os céus e a terra.(1)

A cultura portuguesa tem raízes na cultura judaico-cristã. E é principalmente ao nível da linguagem religiosa que vão encontrar-se mais termos de origem hebraica, registados em todos os vocabulários e dicionários da língua portuguesa.

> Amém, belzebu, cabala, éden, hossana, rabi, sábado, torá, etc.

Nas traduções dos textos bíblicos para português também podem encontrar-se os nomes das letras do **alfabeto hebraico** ou dos meses do **calendário judaico**.

> A primeira tradução do *Antigo Testamento* para português data de 1694 e deve-se aos pastores protestantes João Ferreira de Almeida e Jacobus op den Akker, tendo sido publicada em vários volumes e edições, ao longo da primeira metade do século XVIII, na colónia dinamarquesa de Tranquebar (na atual Índia) e na colónia holandesa de Batávia (ilha de Java, na atual Indonésia)(2). Este interesse dos colonizadores protestantes revela o papel de língua franca que o português desempenhava à época no Índico oriental. Recorda-se que ainda no século XIX os tratados entre o Reino do Sião e outros países eram redigidos nas línguas das partes e em língua portuguesa.
>
> A designada **Bíblia de Almeida** é a base de muitas edições utilizadas pelas igrejas protestantes lusófonas(3). As editoras católicas portuguesas utilizam outras traduções, como a Bíblia de Jerusalém, a Bíblia Pastoral ou a Bíblia para Todos. Note-se que o Vaticano, embora o faça noutras línguas, não disponibiliza uma versão da Bíblia em português(4). As Testemunhas de Jeová utilizam uma tradução própria, a Bíblia do Novo Mundo, disponibilizada na Internet(5).

Nas traduções editadas em Portugal, porém, é quase uma questão de fé encontrar os nomes dessas letras e meses em ortografia portuguesa, pois na maioria empregam-se transliterações várias, estrangeiras às regras do português, talvez consideradas mais eruditas ou «técnicas». Felizmente, no espaço da lusofonia há sempre tradutores bíblicos que usam nomes em ortografia portuguesa, tornando talvez os textos mais acessíveis para um maior número de leitores de português e ajudando, simultaneamente, a **fixar a pronúncia**.

Assim, pode destacar-se, pela coerência da utilização dos aportuguesamentos, a biblioteca digital da Torre de Vigia(6). Destacam-se também, para os nomes das letras, a edição católica da Bíblia para Todos da Paulinas Editora em colaboração com a Sociedade Bíblica e, para os nomes dos meses, a edição pastoral da Paulus Editora ou a versão corrigida e atualizada da Bíblia de Almeida disponibilizada pela União Sefardita Hispano-Portuguesa(7). Soluções em português podem também ser encontradas noutras edições em linha, como a Bíblia Ave Maria(8) ou outras versões eletrónicas(9).

Note-se que o AO90 tem disposições específicas para os textos bíblicos.

> 4 — Os dígrafos finais de origem hebraica *ch*, *ph* e *th* podem conservar-se em formas onomásticas da tradição bíblica, como *Baruch*, *Loth*, *Moloch*, *Ziph*, ou então simplificar-se: *Baruc*, *Lot*, *Moloc*, *Zif*. Se qualquer um destes dígrafos, em formas do mesmo tipo, é invariavelmente mudo, elimina-se: *José*, *Nazaré*, em vez de *Joseph*, *Nazareth*; e se algum deles, por força do uso, permite adaptação, substitui-se, recebendo uma adição vocálica: *Judite*, em vez de *Judith*.
> 5 — As consoantes finais grafadas *b*, *c*, *d*, *g* e *t*(10) mantêm-se, quer sejam mudas quer proferidas, nas formas onomásticas em que o uso as consagrou, nomeadamente antropónimos/antropônimos e topónimos/topônimos da tradição bíblica; *Jacob*, *Job*, *Moab*, *Isaac*, *David*, *Gad*; *Gog*, *Magog*;

Bensabat, *Josafat*. (…) Nada impede, entretanto, que dos antropónimos/antropônimos em apreço sejam usados sem a consoante final Jó, Davi e Jacó.[(11)]

Letras hebraicas

O consonantário ou escrita consonântica é um sistema de escrita que utiliza um alfabeto apenas com consoantes — é o caso do hebraico e do árabe. O alfabeto hebraico é composto por 22 consoantes de base. Cinco consoantes têm uma forma diferente quando ocorrem no final das palavras. Algumas consoantes podem representar mais do que um som, quando modificadas pela adição de um diacrítico, o daguexe — ponto de geminação.

Ex.: בּ (com daguexe: /b/), ב (sem daguexe: /v/).

Para auxiliar a leitura — que se faz da direita para a esquerda —, podem ser utilizados vários sinais diacríticos para representar as vogais, os nicudes — sinais massoréticos.

Ex.: ת (sem nicude: /t/), תָ (com um nicude: /tä/).

Segue-se tabela com os nomes em português das letras do alfabeto hebraico tal como encontrados em várias edições bíblicas, vocabulários e dicionários em língua portuguesa. Registam-se apenas os nomes que obedecem às regras ortográficas do português.

	normal/final	fonética	he	pt	en	fr
1	א	/ʔ/	אָלֶף	**álefe** [TV, BT, Mich, VOLPB, VOLPC] alefe [Mich]	aleph alef	aleph
2	בּ ב	/b/ /v/	בֵּית	**bete** [TV, BT, AD, VOP, VOLPB]	beth bet beith	beth bèt
3	גּ ג	/g/ /ɣ/	גִּימֶל	**guímel** [TV, AD, Inf, VOP, VOLPB, VOLPC] guimel [BT]	gimel	ghimel gimel guimel
4	דּ ד	/d/ /ð/	דָּלֶת	**dálete** [TV, BT, VOP, VOLPB] dalete [AD]	daleth dalet dallet	daleth dalet dalèt
5	ה	/h/	הֵא	**hê** [TV, BT, VOP, VOLPC] hé [AD]	he hei	hé hè he
6	ו	/v/	וָו	**vau** [TV, BT, VOLPB, VOLPC] uau [AD, VOP]	waw vav	vau vav waw
7	ז	/z/	זַיִן	**zaine** [TV, VOP, VOLPB]	zayin zain	zaïn zayin
8	ח	/x/	חֵית	**hete** [TV, BT, VOP, VOLPB]	heth het heith	heth cheth het hèt
9	ט	/t/	טֵית	**tete** [TV, BT, Inf, VOP]	teth tet teith	teth tet tèt
10	י	/j/ ou /y/	יוֹד	**iode** [TV, AD, Inf, Prib, VOP, VOLPB, VOLPC] jode [BT]	yodh yod	iod yod youd
11	כּ / ךּ כ / ך	/k/ /χ/ ou /x/	כַּף	**cafe** [TV, BT]	kaph kaf khaf	caph

12	ל	/l/	לָמֶד	**lâmede** [TV, BT, AD, VOLPB]	lamedh lamed lammed	lamed
13	מ / ם	/m/	מֵם	**meme** [TV, AD, VOP, VOLPB] mem [BT]	mem	mem
14	נ / ן	/n/	נוּן	**nune** [TV, AD, VOP, VOLPB] num [BT]	nun nuun	nun
15	ס	/s/	סָמֶךְ	**sâmeque** [TV, BT, VOP] sameque [AD]	samekh	samech
16	ע	/ħ/	עַיִן	**aine** [TV, VOP, VOLPB]	ayin ayn	hgain aïn oen
17	פּ / ףּ פ / ף	/p/ /f/	פֵּא	**pê** [TV, BT, VOP, VOLPB, VOLPC]	pe peh	pé phé
18	צ / ץ	/ts/	צָדִי	**tsadê** [TV, BT]	tsade tsadde tsadi	sadi tsadé
19	ק	/k/	קוֹף	**cofe** [TV, BT, VOP]	qoph qop khof qof	koph
20	ר	/ʁ/	רֵישׁ	**rexe** [TV, AD, VOP, VOLPB] reche [BT]	resh reish	resch
21	שׁ שׂ	/ʃ/ /s/	שִׁין	**xine** [AD, VOP, VOLPB] chim [TV] sim [TV]	shin	sin schin
22	תּ ת	/t/ /θ/ ou /s/	תָּו	**tau** [TV, BT, Inf, VOP, VOLPB, VOLPC]	taw tav	tau

Fontes: Torre de Vigia (TV); Bíblia Pastoral — Paulus (BP); Bíblia de Jerusalém — Paulus (BJ); Bíblia para Todos — Paulinas (BT); Bíblia dos Capuchinhos — Paulinas (BC); Bíblia Ave Maria (BAM); Bíblia de Almeida Revista e Corrigida (ARC); Bíblia de Almeida Revista e Atualizada (ARA); Aulete Digital (AD); Infopédia (Inf); Priberam (Prib); Michaelis (Mich); Vocabulário Ortográfico Português (VOP); Vocabulário Ortográfico da Língua Portuguesa da Academia Brasileira de Letras (VOLPB); Vocabulário Ortográfico da Língua Portuguesa da Academia das Ciências de Lisboa (VOLPC)

Meses judaicos

O calendário judaico é lunar, com 12 meses, iniciados em cada lua nova. Adiciona-se um mês suplementar sete vezes cada 19 anos, para compensar a diferença para o ano solar(12). A ordem dos meses no atual calendário judaico difere da ordem indicada na Bíblia, correspondente ao calendário anterior ao exílio do povo hebreu(13) na Babilónia (calendário pré-exílico), tendo também os nomes de alguns meses mudado após o exílio(14). Em Israel é utilizado igualmente o calendário gregoriano.

Segue-se tabela com os nomes em português dos meses do atual calendário judaico tal como encontrados em várias edições bíblicas, vocabulários e dicionários em língua portuguesa. Registam-se apenas os nomes que obedecem às regras ortográficas do português, adaptados ao Acordo Ortográfico de 1990 (nomes dos meses com minúscula inicial).

Bíblia	atual	he	pt	en	fr
7.º	1.º	תִּשְׁרֵי	**tisri** [TV, Inf]	Tishri Tishrei	tichri
8.º	2.º	מַרְחֶשְׁוָן מרחשוון	**chesvã** [TV]	Heshvan Marẖeshvan Marcheshvan Cheshvan Marẖeshwan	hèchvan

9.º	3.º	כִּסְלֵו כסליו	**quisleu** [TV, BT, ARC, BC] casleu [BAM] chisleu [BT]	Kislev Kislev Chisleu Chislev	kislev
10.º	4.º	טֵבֵת	**tebete** [TV] tebet [BJ, BT, BP, BAM] tébete [Aur] tébet [Inf, BC]	Tebet Tevet Tebeth	téveth
11.º	5.º	שְׁבָט	**sebate** [TV] sebat [ARC, BC] sabat [BJ, BAM] chebat [BT, Inf, BC]	Shebat Shvat Shevat Sebat	chevat
12.º	6.º	אֲדָר א׳	**adar** [TV, BJ, VOLPB, AD, Mich, BT, Inf, ARC, BC; BAM]	Adar I	adar I
supl.	supl.	אֲדָר אֲדָר ב׳	**adar II** veadar [TV]	Adar II	adar II
1.º	7.º	נִיסָן	**nisã** [TV, BP, BJ, BAM, VOLPB, AD, Mich]	Nisan Nissan	nissan
2.º	8.º	אִיָּר אייר	**íiar** [TV] iiar [BJa] iar [BT, Inf]	Iyar Iyyar	iyar
3.º	9.º	סִיוָן סיוון	**sivã** [TV, BP, BJ, BAM, VOLPBB, Mich]	Sivan Siwan	sivan
4.º	10.º	תַּמּוּז	**tamuz** [TV, BJ, VOLPB, Inf, Mich]	Tammuz Tamuz	tamouz
5.º	11.º	אָב	**ab** [TV, BJ, BT, Inf]	Ab Av	av
6.º	12.º	אֱלוּל	**elul** [TV, BT, BJ, VOLPB, AD, Mich, Inf, ARC, BC, BAM]	Elul	éloul

Ao leitor interessado

O leitor interessado em verificar a ortografia utilizada para as letras e os meses num exemplar da Bíblia deverá consultar:

Esdras	6:15
Neemias	1:1, 2:1, 6:15
Ester	2:16, 3:7, 8:9
Lamentações	1 a 4
Zacarias	1:7, 7:1

O mistério das séries quebradas

Por que razão, nos vocabulários e dicionários em língua portuguesa, aparece a maioria dos nomes das letras do alfabeto hebraico, mas há letras em falta, o mesmo se passando com os meses do calendário judaico? Caso semelhante é o do alfabeto árabe ou dos nomes dos meses do calendário da Hégira[15].

Por que razão:

- o Aulete Digital[16] tem as letras bete (2.ª), guímel (3.ª), dalete (4.ª), hé (5.ª), uau (6.ª), iode (10.ª), lâmede (12.ª), meme (13.ª), nune (14.ª), sameque (15.ª), rexe (20.ª), xine (21.ª), mas não tem as restantes?
- a Infopédia[17] tem quatro letras — guímel (3.ª), tete (9.ª) , iode (10.ª), tau (22.ª) — e não outras?
- o Priberam[18] tem apenas a letra iode (10.ª) e mais nenhuma outra?
- o Michaelis tem apenas a letra álefe (1.ª)?

Já os vocabulários têm mais letras, mas nunca todas:

- o *Vocabulário Ortográfico Português*(19) tem muitas das letras do alfabeto hebraico, mas falta-lhe: álefe, cafe, lâmede e tsadê.
- o *Vocabulário Ortográfico da Língua Portuguesa* da Academia Brasileira de Letras(20) tem muitas das letras do alfabeto hebraico, mas falta-lhe: hê, tete, cafe, sâmeque, tsadê, cofe.
- o *Vocabulário Ortográfico da Língua Portuguesa* da Academia das Ciências de Lisboa(21) não tem: bete, dálete, zaine, hete, tete, cafe, lâmede, meme, nune, sâmeque, aine, tsadê, cofe, rexe, xine.

O Aurélio tem metade dos meses (não tem íiar, ab, chesvã, quisleu, sebate, adar). O Aulete tem três meses (nisã, elul, adar) e o Michaelis cinco meses com grafia portuguesa (nisã, sivã, tamuz, elul, adar). O *Vocabulário Ortográfico da Língua Portuguesa* da Academia Brasileira de Letras tem cinco meses (nisã, sivã, tamuz, elul, adar), enquanto o *Vocabulário Ortográfico da Língua Portuguesa* da Academia das Ciências de Lisboa e o *Vocabulário Ortográfico Português* têm zero meses.

O facto de apenas fontes brasileiras apresentarem os nomes dos meses judaicos em ortografia do português leva a que se possa questionar, por paralelismo com o caso do calendário da Hégira (em que ramadã no Brasil é ramadão em Portugal), se nisã, sivã e chesvã não deveriam ser nisão, sivão e chesvão em Portugal.

Paulo.Correia@ec.europa.eu
Jorge-Madeira.Mendes@ec.europa.eu

---

(1) União Sefardita Hispano-Portuguesa, *Bíblia — Versão Hebraica do Mekhon_Mamre*, «Génesis», http://www.judaismo-iberico.org/interlinear/tanakh/0101PT.HTM.
(2) *A Sentinela — 2007*, 1.º de julho, «A primeira Bíblia em português — uma história de perseverança», https://wol.jw.org/pt/wol/d/r5/lp-t/2007483.
(3) A Bíblia completa em um único volume só foi publicada em 1819. A edição de 1898, feita na Europa, viria a ser conhecida como «Revista e Corrigida». Em meados do século XX, no Brasil, o texto de Almeida foi revisto e atualizado e essa edição é conhecida como «Revista e Atualizada». Sociedade Bíblica do Brasil, *A Bíblia Sagrada: João Ferreira de Almeida*, http://www.sbb.org.br/a-biblia-sagrada/joao-ferreira-de-almeida/.
(4) Vaticano, *Sagrada Escritura*, http://www.vatican.va/archive/bible/index_po.htm.
(5) Bíblias em Português: Tradução do Novo Mundo da Bíblia Sagrada, https://wol.jw.org/pt/wol/binav/r5/lp-t.
(6) Biblioteca *on-line* da Torre de Vigia, https://wol.jw.org/pt/wol/h/r5/lp-t.
(7) União Sefardita Hispano-Portuguesa, *Bíblia — Versão Hebraica do Mekhon_Mamre*, http://www.judaismo-iberico.org/interlinear/tanakh/indexpt.htm.
(8) Bíblia Católica, *Bíblia Ave Maria*, «Génesis», http://www.bibliacatolica.com.br/biblia-ave-maria/.
(9) Bíblia On Line, *Pesquisa Avançada*, http://www.bibliaonline.net/acessar.cgi?pagina=avancada&lang=pt-BR.
(10) No original está «h». Supõe-se que se queria dizer «t», dado que os exemplos apresentados são Bensabat e Josafat.
(11) Portal da Língua Portuguesa, *Acordo Ortográfico da Língua Portuguesa de 1990*, http://www.portaldalinguaportuguesa.org/acordo.php?action=acordo&version=1990.
(12) Contrariamente ao calendário muçulmano, em que não são feitas correções.
(13) O povo hebreu dividia-se em 12 tribos, cada uma delas descendente de um dos filhos de Jacob. Depois do exílio na Babilónia, dez das tribos diluíram-se com povos vizinhos, ficando os hebreus reduzidos às tribos de Judá (cujos membros se designavam *judeus*) e Benjamim (minoritária). Portanto, os *hebreus* ficaram na prática reduzidos aos *judeus*, tornando-se estes dois termos quase sinónimos.
(14) Aos meses pré-exílicos de abibe, zive, etanim e bul correspondem os atuais meses de nisã, íiar, tisri e chesvã.
(15) Cf. «Calendário da Hégira e algumas curiosidades», *in* «a folha», n.º 49 — outono de 2015, http://ec.europa.eu/translation/portuguese/magazine/documents/folha49_pt.pdf.
(16) "letra do alfabeto hebraico" site:aulete.com.br
(17) "letra do alfabeto hebraico" site:infopedia.pt
(18) "letra do alfabeto hebraico" site:priberam.pt
(19) Portal da Língua Portuguesa, *Vocabulário Ortográfico Português*, http://www.portaldalinguaportuguesa.org/vop.html.
(20) Academia Brasileira de Letras, *Vocabulário Ortográfico da Língua Portuguesa*, http://www.academia.org.br/nossa-lingua/busca-no-vocabulario.
(21) Academia das Ciências de Lisboa, *Vocabulário Ortográfico da Língua Portuguesa*, http://volp.acad-ciencias.pt/.

Eslovénia — ficha de país

Helena Veiga de Oliveira
Antiga tradutora da Direção-Geral da Tradução — Comissão Europeia
Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

Nesta ficha de país reúne-se informação terminológica relativa à Eslovénia que se encontra dispersa por vários documentos normativos ou de referência das instituições europeias.

Apresenta-se em anexo a esta ficha uma tabela com o alfabeto esloveno, os símbolos fonéticos correspondentes às letras e transliterações aproximadas em português.

REPÚBLICA DA ESLOVÉNIA (IATE: 859299)

CAPITAL: Liubliana
GENTÍLICO/ADJETIVO: esloveno/a(s)
MOEDA: euro
SUBDIVISÃO: cent (forma obrigatória nos atos da UE e a preferir nos demais textos da UE); cêntimo (variante nacional de uso corrente em Portugal e que pode ser utilizada noutro tipo de textos)

Principais cidades:	Liubliana, Maribor, Celje, Kranj
Rios:	Mura, Drava, Sava, Savínia
Serras:	Triglav(1) (nos Alpes Julianos(2))

Subdivisões administrativas

	esloveno	português	inglês	IATE
12	statistična regija	região estatística	statistical region	—
58	upravna enota	unidade administrativa	administrative unit	3529236
212	občina	município	municipality	3553047

Fonte: Eurostat, *Nomenclature of territorial units for statistics*: National Structures (EU), http://ec.europa.eu/eurostat/web/nuts/national-structures-eu.

Regiões

NUTS	SLOVENIJA	ESLOVÉNIA	SLOVENIA	IATE
SI03	**Vzhodna Slovenija**	**Eslovénia Oriental**	**East Slovenia**	**3572844**
SI031	Pomurska	Mura	Pomurska	—
SI032	Podravska	Drava	Podravska	—
SI033	Koroška	Caríntia	Koroška	—
SI034	Savinjska	Savínia	Savinjska	—
SI035	Zasavska(3)	Médio Sava	Zasavska	—
SI036	Posavska(4)	Baixo Sava	Posavska	—
SI037	Jugovzhodna Slovenija	Eslovénia Sudeste	South-East Slovenia	—
SI038	Primorsko-notranjska(5)	Litoral(6)-Carníola Interior	Primorsko-notranjska	—
SI04	**Zahodna Slovenija**	**Eslovénia Ocidental**	**West Slovenia**	**3572845**
SI041	Osrednjeslovenska	Eslovénia Central	Central Slovenia	—
SI042	Gorenjska(7)	Alta Carníola	Gorenjska	—
SI043	Goriška	Gorízia	Goriška	—
SI044	Obalno-kraška	Litoral-Carso(8)	Obalno-kraška	—

Fonte: Serviço das Publicações, *Código de Redação Interinstitucional*: Anexo 10 — Lista das Regiões, http://publications.europa.eu/code/pt/pt-5001000.htm.
Regulamento (CE) n.º 1059/2003 do Parlamento Europeu e do Conselho, de 26 de maio de 2003, relativo à instituição de uma Nomenclatura Comum das Unidades Territoriais Estatísticas (NUTS) (versão consolidada de 2016), http://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:02003R1059-20161219.

Órgãos judiciais

	esloveno	**português**	**inglês**	**IATE**
3	Delovno sodišče	tribunal de trabalho	labour court	3524590
44	Okrajno sodišče	tribunal de comarca	local court; district court(9)	3531538
11	Okrožno sodišče	tribunal de distrito	district court; regional court	913009
1	Upravno sodišče	Tribunal Administrativo	Administrative Court	3524593
1	Ustavno sodišče	Tribunal Constitucional	Constitutional Court	319321
4	Višje sodišče	tribunal superior	higher court	912993
1	Vrhovno sodišče	Supremo Tribunal	Supreme Court	913006

Portal Europeu da Justiça, *Sistemas judiciais nos Estados-Membros — Eslovénia*, https://e-justice.europa.eu/content_judicial_systems_in_member_states-16-si-pt.do?member=1.

Helena.Veiga@hotmail.com
Paulo.Correia@ec.europa.eu

Anexo: Alfabeto esloveno

O alfabeto esloveno é derivado do alfabeto latino, incluindo 25 carateres, três dos quais por adição de diacríticos:

A, B, C, Č, D, E, F, G, H, I, J, K, L, M, N, O, P, R, S, Š, T, U, V, Z, Ž

As letras Q, W, X e Y são utilizadas apenas em palavras estrangeiras, sobretudo nomes.

letra eslovena	**fonética (AFI)**	**equivalente português**	**nome esloveno**	**«transliteração»**
A a	/a/	a (em alto)	**a**pno (cal)	apno
B b	/b/	b (em bola)	**b**arva (cor)	barva
C c	/ts/	ts (em *tsé-tsé*)	**C**elje	Tselie
Č č	/tʃ/	tch (em *Tch*aikovsky)	**č**aj (chá)	tchai
D d	/d/	d (em dado)	**d**an (dia)	dan
E e	/ɛ/ /e/ /ə/	e (em anel) e (em pelo) e (em menino)	km**e**t (agricultor) t**e**den (semana) p**e**s (cão)	kmet teden pes
F f	/f/	f (em faca)	**f**ant (rapaz)	fant
G g	/g/	g (em agora)	**g**oba (cogumelo)	goba
H h	/x/	rr lisboeta (em rosa)	**h**iša (casa)	rixa
I i	/i/	i (em esquina)	**i**gra (jogo)	igra
J j	/j/	i (em cuidado)	**j**ug (sul)	iug
K k	/k/	c (em casa)	**k**ri (sangue)	kri
L l	/l/ /w/	l (em lua) u (em mau)	**l**ep (bonito) be**l** (branco)	lep beu
M m	/m/	m (em morte)	**m**esto (cidade)	mesto
N n	/n/	n (em nada)	**n**ebo (céu)	nebo
O o	/ɔ/ /o/	o (em morte) o (em morto)	v**o**da (água) m**o**rje (mar)	voda morie
P p	/p/	p (em pato)	**p**et (cinco)	pet
R r	/r/	r (em caro)	**r**oka (mão)	roka
S s	/s/	s (em sapo)	**s**in (filho)	sin
Š š	/ʃ/	x (em xá)	**š**est (seis)	xest
T t	/t/	t (em tudo)	**t**elo (corpo)	telo
U u	/u/	u (em peru)	**u**lica (rua)	ulitsa
V v	/v/ /w/	v (em ave) u (em quase)	**v**olk (lobo) pra**v**ljica (conto de fadas)	volk praulitsa
Z z	/z/	z (em zebra)	**z**id (parede)	zid
Ž ž	/ʒ/	j (em jeito)	**ž**enska (mulher)	jenska

---

(1) Montanha representada nas armas da Eslovénia. Triglav — literalmente, «três cabeças». Tricorno, em italiano; Terglau, em alemão.

(2) Em sentido amplo, os Alpes Julianos incluem igualmente a cadeia de Caravancas e os Alpes de Kamnik-Savínia.

(3) *Za-*, prefixo que significa algo como «além-». Médio Sava indica a posição da região em relação ao curso esloveno do rio Sava.

(4) *Po-*, prefixo que significa algo como «riba-». Baixo Sava indica a posição da região em relação ao curso esloveno do rio Sava.

(5) Até 1 de janeiro de 2015, Notranjsko-kraška.

(6) Primorska, «Litoral», embora a atual região Litoral-Carníola Interior não tenha acesso ao mar.

(7) Kranjska — Carníola.

(8) Kras (esloveno), Carso (italiano), Karst (alemão). Deu origem ao adjetivo **cársico**, aplicado ao tipo de relevo resultante da ação dissolvente das águas (superficiais e subterrâneas) característico de regiões com espessas camadas calcárias.

(9) Contrariamente à terminologia adotada na versão inglesa do Portal Europeu da Justiça («*Courts with general jurisdiction include 44 local, 11 district, and 4 higher courts and the Supreme Court*»), fontes eslovenas em língua inglesa indicam que os tribunais com competências genéricas incluem «*44 district, 11 regional, and 4 higher courts, and the Supreme Court*». Cf. Slovenia.si, *State*, «The Judicial System», http://www.slovenia.si/slovenia/state/the-judicial-system/.

Redação: Paulo Correia (Comissão); Valdemar Ferreira (PE); Manuel Leal (Conselho da UE); Victor Macedo (CESE-CR); António Raúl Reis (Serviço das Publicações)
Grupo de apoio: Ana Luísa Faria (Conselho da UE); Hilário Leal Fontes (Comissão); Susana Gonçalves (Comissão); Ana Lorenzo Garrido (Comissão); Joana Seixas (CESE-CR)
Paginação: Susana Gonçalves (Comissão)
Envio de correspondência: dgt-folha@ec.europa.eu

Edição impressa: oficinas gráficas do Serviço de Infraestruturas e Logística — Bruxelas (Comissão)
Edição eletrónica: sítio Web da Direção-Geral da Tradução da Comissão Europeia no portal da União Europeia — http://ec.europa.eu/translation/portuguese/magazine

Centro de Informação Europeia Jacques Delors
DGAE/Ministério dos Negócios Estrangeiros

https://infoeuropa.eurocid.pt/opac/?func=find-c&ccl_term=wti=(a+folha)

«a folha» ISSN 1830-7809

ISSN 1830-780-9
9 771830 780004